

STIC-Biotech/ChemLib

140214

From: Chan, Christina
Sent: Monday, December 13, 2004 4:29 PM
To: Basi, Nirmal; STIC-Biotech/ChemLib
Subject: RE: Rush search for 09/995,225

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Basi, Nirmal
Sent: Monday, December 13, 2004 1:49 PM
To: Chan, Christina
Subject: Rush search for 09/995,225

Examiner: Nirmal S. Basi
Art Unit 1646
Office: Remsen Building, Room 4D68
Mail Room: Remsen Building, room 4C70

Sequence search:

App. #: 09/995,225
Result format: Paper.

Title: ENDOGENOUS AND NON-ENDOGENOUS VERSIONS OF HUMAN G PROTEIN-
COUPLED RECEPTORS

Inventors: Ruoping Chen et al

Priority Date: 10/13/98

Please search:

- i) Please search SEQ ID NO: 1, 2, 15, 16, 41 and 42,
- ii) Polynucleotide encoding the polypeptide of SEQ ID NO: 16

Search issued and commercial databases.

RECEIVED
DEC 13 2005
STIC

Thanks,

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: 12/14/04
Date Completed: 12/15/04
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search 4
NA Sequence: # _____
AA Sequence: # 271
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: 9601
WWW/Internet: _____
Other(Specify): _____

revert to
NA

This Page Blank (uspto)

QY	1	ATGAGACAAACGACGCCCCACTCTCGAGCCAAACAGTCCGTGTCTTGAGTGTCCTCCCGGC	60
Db	1	ATGAGACAAACGACGCCCCACTCTCGAGCCAAACAGTCCGTGTCTTGAGTGTCCTCCCGGC	60
QY	61	TGGGCGTGGGGCTTGGGGTTTCCGACCAGTGATCTAACAGCCCTCTTGCGTGAGCCCGGT	120
Db	61	TGGGCGTGGGGCTTGGGGTTTCCGACCAGTGATCTAACAGCCCTCTTGCGTGAGCCCGGT	120
QY	121	TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAACAGCAAG	180
Db	121	TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAACAGCAAG	180
QY	181	TCCTCCTCAACATATCTCTTGAGCACTCGGTGCTGGCCGGAATCTTGCTCTCTTTTCATA	240
Db	181	TCCTCCTCAACATATCTCTTGAGCACTCGGTGCTGGCCGGAATCTTGCTCTCTTTTCATA	240
QY	241	GTGTTGTGGACCTTCCTGTGGAAAGATTTCATCTTGAAATGACAGATGCTCAGAGTCCC	300
Db	241	GTGTTGTGGACCTTCCTGTGGAAAGATTTCATCTTGAAATGACAGATGCTCAGAGTCCC	300
QY	301	GACCAAGATCATAGAAATGTGTGAATTCATCTATCTCAACACTTCAATAGATATCTGTA	360
Db	301	GACCAAGATCATAGAAATGTGTGAATTCATCTATCTCAACACTTCAATAGATATCTGTA	360
QY	361	CCGTTAACCATGACAGAGTATATCGCTGTCTGCCAACCCGCTCAAGTACACACGCGTCA	420
Db	361	CCGTTAACCATGACAGAGTATATCGCTGTCTGCCAACCCGCTCAAGTACACACGCGTCA	420
QY	421	TACCAAGCCCGCACCCCGGAAAGTCACTGTAACTGTTTACATCACTGCTTCTGACACG	480
Db	421	TACCAAGCCCGCACCCCGGAAAGTCACTGTAACTGTTTACATCACTGCTTCTGACACG	480
QY	481	ATCCCTATTAATCTGGTGGCCCAACATCTGGAAGCTGAATCAATGACACCTCTGAT	540
Db	481	ATCCCTATTAATCTGGTGGCCCAACATCTGGAAGCTGAATCAATGACACCTCTGAT	540
QY	541	CACGTCCTCATCTGGATCCACTGCTTCAACCGATCTAACCTGGTCCCTGCTCAATCTTCTTC	600
Db	541	CACGTCCTCATCTGGATCCACTGCTTCAACCGATCTAACCTGGTCCCTGCTCAATCTTCTTC	600
QY	601	ATCTTGAAGCTCAATCATATGTTGTGAACAAGCTCAGAGAAAGAAAGCAATTTTGTCTCCGTGGC	660
Db	601	ATCTTGAAGCTCAATCATATGTTGTGAACAAGCTCAGAGAAAGAAAGCAATTTTGTCTCCGTGGC	660
QY	661	TACTTCAAGGGGAAAGACACCGCACATCTGTTCACACATTAACCTCCATCTTTGCCACATT	720
Db	661	TACTTCAAGGGGAAAGACACCGCACATCTGTTCACACATTAACCTCCATCTTTGCCACATT	720
QY	721	TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATAGGGGCGCCATCCAGAACCGC	780
Db	721	TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATAGGGGCGCCATCCAGAACCGC	780
QY	781	TGGCTGTGACATCATGATGCCAGCAATTTGGCAACATGACCTTCTGAAACACGCAAT	840
Db	781	TGGCTGTGACATCATGATGCCAGCAATTTGGCAACATGACCTTCTGAAACACGCAAT	840
QY	841	AACCTTCTCTCTATCTGCTTATCAGCAAGCGGTTCCGACCATAGAGACGCGCACGCTC	900
Db	841	AACCTTCTCTCTATCTGCTTATCAGCAAGCGGTTCCGACCATAGAGACGCGCACGCTC	900
QY	901	AAGGTTTCTCAAGTGCAGAGAACCTGTGAACGTTCTACACCAATCATTAATCTTTTC	960
Db	901	AAGGTTTCTCAAGTGCAGAGAACCTGTGAACGTTCTACACCAATCATTAATCTTTTC	960
QY	961	ATTAACAAGTAGCCCTGTGATCTCGCCGCAAACTACACTGATCAAGAATGCTGTGTATC	1020
Db	961	ATTAACAAGTAGCCCTGTGATCTCGCCGCAAACTACACTGATCAAGAATGCTGTGTATC	1020
QY	1021	CAGATATGCAAAAATGAAAAACCTATTAAGAATATCCCGTGA 1062	
Db	1021	CAGATATGCAAAAATGAAAAACCTATTAAGAATATCCCGTGA 1062	

[illegible]

OY	661	TACTCAGGGGAAACACACGGCATCTTGTTAACATTA	720
Db	661	TACTCAGGGGAAACACACGGCATCTTGTTAACATTA	720
OY	721	TGGGCCCCCGCATCATGATATCTTTACGACCTTGA	780
Db	721	TGGGCCCCCGCATCATGATATCTTTACGACCTTGA	780
OY	781	TGGCTGTGCACATATGTCGACATTGGCAACATGTA	840
Db	781	TGGCTGTGCACATATGTCGACATTGGCAACATGTA	840
OY	841	AACTCTTCCTTACTGCTTATGACGAAAGGGTTCCG	900
Db	841	AACTCTTCCTTACTGCTTATGACGAAAGGGTTCCG	900
OY	901	AAGGTTTCTTCAAGTGCAGAGCAACCTGTACAGTT	960
Db	901	AAGGTTTCTTCAAGTGCAGAGCAACCTGTACAGTT	960
OY	961	ATAACAGTAGCCCTTGATCTCGCGGCAAACTCACT	1020
Db	961	ATAACAGTAGCCCTTGATCTCGCGGCAAACTCACT	1020
OY	1021	CAGATGACAAAAATGAAAAACCTATAAAGTATCCCG	1062
Db	1021	CAGATGACAAAAATGAAAAACCTATAAAGTATCCCG	1062

RESULT 3				
BD103852		1062 bp	DNA	linear PAT 27-AUG-2002
LOCUS	BD103852			
DEFINITION	Novel G protein-coupled receptor and its DNA.			
ACCESSION	BD103852			
VERSION	BD103852.1			
KEYWORDS	MO 0194582-A/4.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1062)			
AUTHORS	Terao,Y., Matsui,H. and Shintani,Y.			
TITLE	Novel G protein-coupled receptor and its DNA			
JOURNAL	Patent: WO 0194582-A 4 13-DEC-2001;			
	TAKEDA CHEMICAL INDUSTRIES LTD,YASUKO TERAQ,HIDEKI MATSUI, YASUSHI SHINTANI			
COMMENT	OS Homo sapiens (human)			
	PN MO 0194582-A/4			
	PD 13-DEC-2001			
	PF 01-JUN-2001 WO 2001JP004643			
	PR 02-JUN-2000 JP OOP 170446,23-JUN-2000 JP OOP 194926 PI			
	YASUKO TERAQ,HIDEKI MATSUI,YASUSHI SHINTANI			
	PC C12N5/12,C12P21/02,C07K14/705,C07KL6/28,A61K45/00,A61P25/00,			
	PC A61P29/00,			
	PC A61P9/00,A61P35/00,A61P37/00,A61P1/00,G01N33/566, PC			
	G01N33/50,			
	CC G01N33/15			
	CC Novel G protein-coupled receptor and its DNA			
	FH Key Location/Qualifiers			
	FT source 1..1062			
		/organism='Homo sapiens (human)'. location/Qualifiers 1..1062 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
FEATURES	source			
ORIGIN				
	Query Match	100.0%; Score 1062; DB 6; Length 1062;		
	Beat Local Similarity	100.0%; Prid. No. 5.2e-247;		
	Matches 1062; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
	1 ATGAGCACACGCACGCCCACTCGCAGCCAACAAGCTGGTGTTCGTGTGATCCCCCGAC 60			

Db	1	ATGAGACACACGACGCCCACTCGACGCAACAGCTGGCTGTCTTGGTGGTCCCCCGGC	60
Qy	61	TCGGCTCGGCTTGGGATTTTCGTGCCGTGTCTACTACAGCTCTTGTGCTCGGT	120
Db	61	TCGGCTCGGCTTGGGATTTTCGTGCCGTGTCTACTACAGCTCTTGTGCTCGGT	120
Qy	121	TTACCAAGAAATTTCTTACAGATGATCATCTCTTCCAGCGTGTGTGGCAAGACAGAAAG	180
Db	121	TTACCAAGAAATTTCTTACAGATGATCATCTCTTCCAGCGTGTGTGGCAAGACAGAAAG	180
Qy	181	TCCTCTCAACATCATCTTGTGGCACTGCTGTGCGGACATCTTGTGCTCTTCTTTCATA	240
Db	181	TCCTCTCAACATCATCTTGTGGCACTGCTGTGCGGACATCTTGTGCTCTTCTTTCATA	240
Qy	241	GTTGTTTGGACCTTCTGTGTGGAAGATTTCATTGAACATGACAGATGCTTCAGGTCCCC	300
Db	241	GTTGTTTGGACCTTCTGTGTGGAAGATTTCATTGAACATGACAGATGCTTCAGGTCCCC	300
Qy	301	GACCAAGATCATAGAAAGTCTGGAATTTCTCATCCATCCACACTCTCCATATGATTTACTGTA	360
Db	301	GACCAAGATCATAGAAAGTCTGGAATTTCTCATCCATCCACACTCTCCATATGATTTACTGTA	360
Qy	361	CGGTTAACATTTGACAGGTATATGCGTGTCTGTGCCACCGGCTCAAGTACACACGGTCTCA	420
Db	361	CGGTTAACATTTGACAGGTATATGCGTGTCTGTGCCACCGGCTCAAGTACACACGGTCTCA	420
Qy	421	TACCCAGCCGCGCACCCGGAAAGTCATTGTAAGTGTTTACATCACTGCTTCTGTACAGC	480
Db	421	TACCCAGCCGCGCACCCGGAAAGTCATTGTAAGTGTTTACATCACTGCTTCTGTACAGC	480
Qy	481	ATCCGCTTATTACTGTGTGGCCCAACATCTGGACTGAAGCTACATCAGACCTCTGTGAT	540
Db	481	ATCCGCTTATTACTGTGTGGCCCAACATCTGGACTGAAGCTACATCAGACCTCTGTGAT	540
Qy	541	CACGCTCATCTGGAATCCATGCTGTCAACGATCACTGAGGCGCTGCTCCATCTTCTTC	600
Db	541	CACGCTCATCTGGAATCCATGCTGTCAACGATCACTGAGGCGCTGCTCCATCTTCTTC	600
Qy	601	ATCTTGAACCTCAATCATCTGTTACAGAGCTCAGAGGAAGACAAATTTTCGTCTCGGTGC	660
Db	601	ATCTTGAACCTCAATCATCTGTTACAGAGCTCAGAGGAAGACAAATTTTCGTCTCGGTGC	660
Qy	661	TACTTCCAGGGGGAAGACACCGGCATCTTGTTCACATTAAGCTTCTTGTGCAACACTT	720
Db	661	TACTTCCAGGGGGAAGACACCGGCATCTTGTTCACATTAAGCTTCTTGTGCAACACTT	720
Qy	721	TGGGCCCCCGGCATCATGATTTCTTTTACACACTCTATGGGGGCGCCATCCAGAACCGC	780
Db	721	TGGGCCCCCGGCATCATGATTTCTTTTACACACTCTATGGGGGCGCCATCCAGAACCGC	780
Qy	781	TGGCTGTGACATCATGTGCCAGATTGGCAACATGCTAGGCTTCTGTGAACAGACCATC	840
Db	781	TGGCTGTGACATCATGTGCCAGATTGGCAACATGCTAGGCTTCTGTGAACAGACCATC	840
Qy	841	AACCTTCTTCTCTACTGTCTTATCAGCAAGCGGTTTCGCAACATAGGCGCGCCACGCTC	900
Db	841	AACCTTCTTCTCTACTGTCTTATCAGCAAGCGGTTTCGCAACATAGGCGCGCCACGCTC	900
Qy	901	AAGGCTTCTTCAAGTGCAGAAACAACCTGTACAGTTCTTACACCAATCATTAATTTTCC	960
Db	901	AAGGCTTCTTCAAGTGCAGAAACAACCTGTACAGTTCTTACACCAATCATTAATTTTCC	960
Qy	961	ATTAACAAGTAGCCCTGTGATCTGCGCGGCAACTCACTGACATCAAGATCTGTGTAC	1020
Db	961	ATTAACAAGTAGCCCTGTGATCTGCGCGGCAACTCACTGACATCAAGATCTGTGTAC	1020
Qy	1021	CAGATATGACAAATATGAAAAACCTATAAAGTATCCCCCGTGA	1062
Db	1021	CAGATATGACAAATATGAAAAACCTATAAAGTATCCCCCGTGA	1062

RESULT 4

AY635179 1167 bp mRNA linear PRI 04-JUN-2004
LOCUS
DEFINITION Homo sapiens G protein-coupled receptor 139 (GPR139) mRNA, complete
ACCESSION AY635179
VERSION AY635179.1 GI:49413551
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 1167)
TITLE Bonner, T.I., Nagle, J.W. and Kauffman, D.
JOURNAL Complete coding sequence of GPR139
REFERENCE 2 (bases 1 to 1167)
AUTHORS Bonner, T.I., Nagle, J.W. and Kauffman, D.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2004) Lab of Genetics, NIMH, Bldg 36, Rm 3D06,
MSC4094, Bethesda, MD 20892-4094, USA
FEATURES
source
1..1167
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p12.3"
/issue_type="brain"
1..1167
/gene="GPR139"
92..1153
/gene="GPR139"
/codon_start=1
/product="G protein-coupled receptor 139"
/protein_id="AA05818.1"
/db_xref="GI:49413552"
/translation="MEHTRAHLAANSSLSWSPGACGLGVPPVYVYSLILCLGHPAN
ILTVIILSLVARQKSSIVYLLALAAADILVFLVDFLLEDFILMOMQVDPK
ILVLESSLHSISIMVLPITLDYIAVCHPKMYHSPATRAKIVYVYTCFLTS
IPYWPMPNIMTBDYISTVHVLIMHCFVILVPCSPFINSIIVYLRKSNRL
RGYSTGKTALIFTITSIPATLMAPRIIMILVHVGAPIONNIVHISMDIANMALL
KNALNPFVLCNKSPFRMAAATLAKAFKFCQKQPOVPTNHFSTSSFWISFANSHC
IKMLVYQYKSGKPKIVSP"
complement (1142..1167)
ORIGIN
primer_bind
Query Match 100.0%; Score 1061.6; DB 9; Length 1167;
Best Local Similarity 99.9%; Pred. No. 6.5e-247;
Matches 1061; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCAGACGACGACCCCACTTCGACGCAACAGCTGCTGTGTTGGTCCCGCGGC 60
DB 92 ATGGAGCAGACGACGACCCCACTTCGACGCAACAGCTGCTGTGTTGGTCCCGCGGC 151
QY 61 TCGGCGCTGGCGCTTGGGTTTCGTGCCCGGTGCTCTACAGCCTCTGCTGCTCCGT 120
DB 152 TCGGCGCTGGCGCTTGGGTTTCGTGCCCGGTGCTCTACAGCCTCTGCTGCTCCGT 211
QY 121 TTACCAAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGAG 180
DB 212 TTACCAAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGAG 271
QY 181 TCCCTCTCAACTATCTCTGGCACTCGCTGTCGCGCAATCTTGCTCTTTTTCATA 240
DB 272 TCCCTCTCAACTATCTCTGGCACTCGCTGTCGCGCAATCTTGCTCTTTTTCATA 331
QY 241 GTGTTTGGGACTTCTGTTGGAGATTTCATCTTGAAACATGACAGATGCCAGAGTCCC 300
DB 332 GTGTTTGGGACTTCTGTTGGAGATTTCATCTTGAAACATGACAGATGCCAGAGTCCC 391
QY 301 GACCAAGATCATAGAGTCTGGAATTCATCATCCATCCACACTTCATATGATTAAGTGA 360
DB 392 GACCAAGATCATAGAGTCTGGAATTCATCATCCATCCACACTTCATATGATTAAGTGA 451

QY 361 CGGTAAACATTGACAGGATATCGGTGTCGACCCGCTGAAGTACACAGGTTCA 420
DB 452 CGGTAAACATTGACAGGATATCGGTGTCGACCCGCTGAAGTACACAGGTTCA 511
QY 421 TACCCAGCCCGACCCGGAAGTCAATTGAAGTTTATCATCACTGCTCTGACAGC 480
DB 512 TACCCAGCCCGACCCGGAAGTCAATTGAAGTTTATCATCACTGCTCTGACAGC 571
QY 481 ATCCCTATTAATGATGGGCGCAACATCTGACATGAAGACTACATGACACTTGTGCAT 540
DB 572 ATCCCTATTAATGATGGGCGCAACATCTGACATGAAGACTACATGACACTTGTGCAT 631
QY 541 CAGTCTCATCTGATCCATCACTGCTTCAACCGTCTACCTGGGCGCTGCTCATCTTTC 600
DB 632 CAGTCTCATCTGATCCATCACTGCTTCAACCGTCTACCTGGGCGCTGCTCATCTTTC 691
QY 601 ATCTTGAATCAATATGTTGTACAGAGTCAAGAGGAGAGCAATTTTCGTCCGTCGAC 660
DB 692 ATCTTGAATCAATATGTTGTACAGAGTCAAGAGGAGAGCAATTTTCGTCCGTCGAC 751
QY 661 TACTCAGCGGGAAGACACCGGCACTTGTTCACAACTTACCTCATCTTTCGACACTT 720
DB 752 TACTCAGCGGGAAGACACCGGCACTTGTTCACAACTTACCTCATCTTTCGACACTT 811
QY 721 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTATGGGGCGCCCATCAGAACCGC 780
DB 812 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTATGGGGCGCCCATCAGAACCGC 871
QY 781 TGGCTGTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 872 TGGCTGTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 931
QY 841 AACTTCTCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 932 AACTTCTCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 991
QY 901 AAGCTTTCTTCAAGTCCGAGAGCAACCTGTACAGTTCTTACCAATCAATCAATCTTTCC 960
DB 992 AAGCTTTCTTCAAGTCCGAGAGCAACCTGTACAGTTCTTACCAATCAATCAATCTTTCC 1051
QY 961 ATTAACAAGTACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1052 ATTAACAAGTACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
QY 1021 CAGTATGACAAAATGGAATCTATPAAAGTATCCCGTGA 1062
DB 1112 CAGTATGACAAAATGGAATCTATPAAAGTATCCCGTGA 1153
RESULT 5
BD183203
LOCUS
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD183203
VERSION BD183203.1 GI:31875403
KEYWORDS JP 2002345481-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Terao, Y., Matsui, H. and Shintani, Y.
TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: JP 2002345481-A 3 03-DEC-2002;
TAKEDA CHEMICAL INDUSTRIES LTD
OS Homo sapiens (human)
PN JP 2002345481-A/3
PD 03-DEC-2002
PF 01-JUN-2001 JP 2001166688
PI YASUOKA TERAOKA HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/09, A61K45/00, A61P1/00, A61P9/00, A61P25/00, PC
A61P29/00,

PC A61P35/00,A61P37/00,C07K14/705,C07K16/28,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/66,G01N33/15,G01N33/50,G01N33/53, PC
G01N33/53,
PC G01N33/566,C12N15/00,C12N5/00
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT Source 1..1062
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

Source

ORIGIN

Query Match 99.8%; Score 1060.4; DB 6; Length 1062;
Best Local Similarity 99.9%; Pred. No. 1.3e-246;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGCCACCTCGACGCCAAGCTCGCTGTTGGTGTCCCGGCG 60
DB 1 ATGAGACACACGACGCCACCTCGACGCCAAGCTCGCTGTTGGTGTCCCGGCG 60
QY 61 TGGGCTGGGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGGTGTCCCGGCG 120
DB 61 TGGGCTGGGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGGTGTCCCGGCG 120
QY 121 TTACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180
DB 121 TTACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180
QY 181 TCTCTCAACAATCTCTTGGCACTCGCTGCTGCGGACATCTTGGTCTTTTCATA 240
DB 181 TCTCTCAACAATCTCTTGGCACTCGCTGCTGCGGACATCTTGGTCTTTTCATA 240
QY 241 GGTGTTGTGACATCTCTTGGCAAGATTTGATCTTGAACATGAGATGCTTCAGTCCC 300
DB 241 GGTGTTGTGACATCTCTTGGCAAGATTTGATCTTGAACATGAGATGCTTCAGTCCC 300
QY 301 GACAAATCATAGAGTGTGAAATCTCATCTCCACACCTCATATGATTTACTGTA 360
DB 301 GACAAATCATAGAGTGTGAAATCTCATCTCCACACCTCATATGATTTACTGTA 360
QY 361 CCGTTAACAATGACAGATATCGCTGTGCGCACCGCTCAAGTAACAACAGGTTCTCA 420
DB 361 CCGTTAACAATGACAGATATCGCTGTGCGCACCGCTCAAGTAACAACAGGTTCTCA 420
QY 421 TACCCAGCCCGCACCCGGAAAGTCAATTTGTAAGTTTACATCACTGCTTCCGACACG 480
DB 421 TACCCAGCCCGCACCCGGAAAGTCAATTTGTAAGTTTACATCACTGCTTCCGACACG 480
QY 481 ATCCCTATTAATGAGTGTGCGCAATCTGGAATGAGACATCAATCAAGCACTCTGTCAT 540
DB 481 ATCCCTATTAATGAGTGTGCGCAATCTGGAATGAGACATCAATCAAGCACTCTGTCAT 540
QY 541 CAGGTCCTCATCTGATTCACCTGCTTCAACGCTTACCTGAGTCCCTGCTCATCTTCTTC 600
DB 541 CAGGTCCTCATCTGATTCACCTGCTTCAACGCTTACCTGAGTCCCTGCTCATCTTCTTC 600
QY 501 ATTCTGAATCATATTTGTAAGCTCAAGAGAGAGAGAAATTTTCTCTCCGTCGC 660
DB 501 ATTCTGAATCATATTTGTAAGCTCAAGAGAGAGAGAAATTTTCTCTCCGTCGC 660
QY 661 TACTCAGCGGAGAGACACCGGCACTTCTTCAACATTAATCTCAATCTTTGCAACACTT 720
DB 661 TACTCAGCGGAGAGACACCGGCACTTCTTCAACATTAATCTCAATCTTTGCAACACTT 720
QY 721 TGGGCTGGGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGGTGTCCCGGCG 780
DB 721 TGGGCTGGGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGGTGTCCCGGCG 780
QY 781 TGGCTGTGACATCATGTGCGACATTTGCCAATGCTACGCTTCTGAAACAGACGATC 840

DB 781 TGGCTGTGACATCATGTGCGACATTTGCCAATGCTACGCTTCTGAAACAGACGATC 840
QY 841 AACTTCTTCTTACTGCTTTCATACGAAAGGTTTCCGACCAATGACCGCCGACGCTC 900
DB 841 AACTTCTTCTTACTGCTTTCATACGAAAGGTTTCCGACCAATGACCGCCGACGCTC 900
QY 901 AAGGCTTCTTCAAGTCCGAAAGCACTGTACAGTTTCAACCAATATTAATCTTTC 960
DB 901 AAGGCTTCTTCAAGTCCGAAAGCACTGTACAGTTTCAACCAATATTAATCTTTC 960
QY 961 ATACAGATGACCGCTGATCTCGCGGCAAACTCACATGATCAAGATGCTGTAC 1020
DB 961 ATACAGATGACCGCTGATCTCGCGGCAAACTCACATGATCAAGATGCTGTAC 1020
QY 1021 CAGTATGACAAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1062
DB 1021 CAGTATGACAAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1062

RESULT 6
BD103851 1062 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD103851
VERSION BD103851.1 GI:22649425
KEYWORDS MO 0194582-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1062)
Terao,Y., Matsumi,H. and Shintani,Y.
Novel G protein-coupled receptor and its DNA
Patent: WO 0194582-A 3 13-DEC-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI
SHINTANI

COMMENT

OS Homo sapiens (human)
PN WO 0194582-A/3
PD 13-DEC-2001
PE 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP 00P 170446,23-JUN-2000 JP 00P 194926 PI
YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/12,C12P21/02,C07K14/705,C07K16/28,A61K45/00,A61P25/00,
PC A61P29/00,
PC A61P37/00,A61P35/00,A61P37/00,A61P1/00,G01N33/566, PC
G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT Source 1..1062
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

Source

ORIGIN

Query Match 99.8%; Score 1060.4; DB 6; Length 1062;
Best Local Similarity 99.9%; Pred. No. 1.3e-246;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGCCACCTCGACGCCAAGCTCGCTGTTGGTGTCCCGGCG 60
DB 1 ATGAGACACACGACGCCACCTCGACGCCAAGCTCGCTGTTGGTGTCCCGGCG 60
QY 61 TGGGCTGGGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGGTGTCCCGGCG 120
DB 61 TGGGCTGGGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGGTGTCCCGGCG 120
QY 121 TTACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180

Db 121 TTACGAGCAATATCTTGAAGATGATCTCTCCAGCTGTGGCAAGAG 180
Qy 181 TCCCTAGACATATCTTGGCACTCGCTGTGCCGACATTTGTCCTTTTTCATA 240
Db 181 TCCCTAGACATATCTTGGCACTCGCTGTGCCGACATTTGTCCTTTTTCATA 240
Qy 241 GTGTTTGAGACTTCTGTGGAGATTTCACTTTGAACATGAGATGCTTCAAGTCCC 300
Db 241 GTGTTTGAGACTTCTGTGGAGATTTCACTTTGAACATGAGATGCTTCAAGTCCC 300
Qy 301 GACAGATCATGAAAGTGTGAATTCATCATCCACCTCCATATGATTAAGTGA 360
Db 301 GACAGATCATGAAAGTGTGAATTCATCATCCACCTCCATATGATTAAGTGA 360
Qy 361 CGTTAACATTTGACAGGATATGCTGTGTGCCACCGCTCAAGTACACAGGCTTCA 420
Db 361 CGTTAACATTTGACAGGATATGCTGTGTGCCACCGCTCAAGTACACAGGCTTCA 420
Qy 421 TACCCAGCCCGACCCGGAAAGTCAATGTAAAGTTTAAATCACTCTGCTTGAACAGC 480
Db 421 TACCCAGCCCGACCCGGAAAGTCAATGTAAAGTTTAAATCACTCTGCTTGAACAGC 480
Qy 481 ATCCCTATTAATGCTGTGCCGCAATCTGGAATGAAGATCAATGACACCTCTGTGCAT 540
Db 481 ATCCCTATTAATGCTGTGCCGCAATCTGGAATGAAGATCAATGACACCTCTGTGCAT 540
Qy 541 CAGTCTCATCTGTGATCACTGCTTCACTGCTTCACTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CAGTCTCATCTGTGATCACTGCTTCACTGCTTCACTGCTGCTGCTGCTGCTGCTGCT 600
Qy 601 ATCTTGAATCAATCATATTTGTGTACAGCTCAAGAGAGAGAAATTTGCTCTCGTGGC 660
Db 601 ATCTTGAATCAATCATATTTGTGTACAGCTCAAGAGAGAGAAATTTGCTCTCGTGGC 660
Qy 661 TACTCCAGGGGAAAGACACCGCATCTGTGACATTAACCTTCACTTTTGCACACTT 720
Db 661 TACTCCAGGGGAAAGACACCGCATCTGTGACATTAACCTTCACTTTTGCACACTT 720
Qy 721 TGGGCCCCCGCATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 780
Db 721 TGGGCCCCCGCATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 780
Qy 781 TGGCTGTGACATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 840
Db 781 TGGCTGTGACATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 840
Qy 841 AACTTTTCTCTTACTGCTTCACTGACAGAGCGGTTCCGACCATGAGCGCGCACGCTC 900
Db 841 AACTTTTCTCTTACTGCTTCACTGACAGAGCGGTTCCGACCATGAGCGCGCACGCTC 900
Qy 901 AAGGCTTTCTTCAAGTGCAGAGACCACTGTGACAGTTTCAACCAATCATTAATTTTCC 960
Db 901 AAGGCTTTCTTCAAGTGCAGAGACCACTGTGACAGTTTCAACCAATCATTAATTTTCC 960
Qy 961 ATPAACAGTAGCCCTGTGATCTGCGCGGAAACTCACTGATCAAGATGCTGTGTAC 1020
Db 961 ATPAACAGTAGCCCTGTGATCTGCGCGGAAACTCACTGATCAAGATGCTGTGTAC 1020
Qy 1021 CAGTATGACAAATAATGAGAAACCTATTAAGATATCCCGTGA 1062
Db 1021 CAGTATGACAAATAATGAGAAACCTATTAAGATATCCCGTGA 1062

RESULT 7
AX254977 1343 bp DNA linear PAT 10-OCT-2001
LOCUS AX254977
DEFINITION Sequence 21 from Patent WO0170978.
ACCESSION AX254977
VERSION AX254977.1 GI:16074504
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS Taupier R.J., Majumder, K., Spaderna, S.K., Smithson, G., Mezes, P.S.
and Vernet, C.A.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0170978-A 21 27-SEP-2001;
Curagen Corporation (US)
FEATURES
source 1..1343
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.7%; Score 1058.8; DB 6; Length 1343;
Best Local Similarity 99.8%; Pred. No. 3.2e-246;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGACACGACGACGACCTCTGACAGCAAGCTGTGTGTTGGTGTCCCGGC 60
Db 2 ATGAGACACGACGACGACCTCTGACAGCAAGCTGTGTGTTGGTGTCCCGGC 61
Qy 61 TGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 120
Db 62 TGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 121
Qy 121 TTACGAGCAAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGGAGAGAGAGAG 180
Db 122 TTACGAGCAAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGGAGAGAGAGAG 181
Qy 181 TCCCTCAACATATCTTGTGGCACTGCTGTGCTGGCAGATCTTGTCTCTTTTTCATA 240
Db 182 TCCCTCAACATATCTTGTGGCACTGCTGTGCTGGCAGATCTTGTCTCTTTTTCATA 241
Qy 241 GTGTTTGAGACTTCTGTGGAAGATTTGATCTTGAACATGAGATGCTTCAAGTCCC 300
Db 242 GTGTTTGAGACTTCTGTGGAAGATTTGATCTTGAACATGAGATGCTTCAAGTCCC 301
Qy 301 GACAGATCATGAAAGTGTGAATTTCTCATTCATCCACACTTCATATGATTAAGTGA 360
Db 302 GACAGATCATGAAAGTGTGAATTTCTCATTCATCCACACTTCATATGATTAAGTGA 361
Qy 361 CCGTTAACATTTGACAGGATATGCTGTGTGCCACCGCTCAAGTACACAGCGTCTCA 420
Db 362 CCGTTAACATTTGACAGGATATGCTGTGTGCCACCGCTCAAGTACACAGCGTCTCA 421
Qy 421 TACCCAGCCCGACCCGGAAAGTCAATGTAAAGTTTAAATCACTCTGCTTGAACAGC 480
Db 422 TACCCAGCCCGACCCGGAAAGTCAATGTAAAGTTTAAATCACTCTGCTTGAACAGC 481
Qy 481 ATCCCTATTAATGCTGTGCCGCAATCTGGAATGAAGATCAATGACACCTCTGTGCAT 540
Db 482 ATCCCTATTAATGCTGTGCCGCAATCTGGAATGAAGATCAATGACACCTCTGTGCAT 541
Qy 541 CAGTCTCATGATGATCACTGCTTCACTGACAGAGCGGTTCCGACCATGAGCGCGCACGCTC 600
Db 542 CAGTCTCATGATGATCACTGCTTCACTGACAGAGCGGTTCCGACCATGAGCGCGCACGCTC 601
Qy 601 ATCTTGAATCAATCATATTTGTGTACAGCTCAAGAGAGAGAAATTTGCTCTCGTGGC 660
Db 602 ATCTTGAATCAATCATATTTGTGTACAGCTCAAGAGAGAGAAATTTGCTCTCGTGGC 661
Qy 661 TACTCCAGGGGAAAGACACCGCATCTTGTGACAGATTAACCTTCACTTTTGCACACTT 720
Db 662 TACTCCAGGGGAAAGACACCGCATCTTGTGACAGATTAACCTTCACTTTTGCACACTT 721
Qy 721 TGGGCCCCCGCATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 780
Db 722 TGGGCCCCCGCATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 781
Qy 781 TGGCTGTGACATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 840
Db 782 TGGCTGTGACATCATGATGATTTTAAACCTTATGAGGCGCCATCCAGAACGCG 841

QY 841 AACTTCTTCTCTACTGCTTATCAGCAAGCGGTTCCGACCATGGAGCGCCGACGCTC 900
 DB 842 AACTTCTTCTCTACTGCTTATCAGCAAGCGGTTCCGACCATGGAGCGCCGACGCTC 901
 QY 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTTACCAATCATTAATTCTTCC 960
 DB 902 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTTACCAATCATTAATTCTTCC 961
 QY 961 ATAAACAAGTACCCCTGGATCTTCCGCGCAACTCACTGATCAAGATGCTGTGTAC 1020
 DB 962 ATAAACAAGTACCCCTGGATCTTCCGCGCAACTCACTGATCAAGATGCTGTGTAC 1021
 QY 1021 CAGTATGACAAAATGAGAAAACCTATTAAGTATCCCGCTGA 1062
 DB 1022 CAGTATGACAAAATGAGAAAACCTATTAAGTATCCCGCTGA 1063
 RESULT 8
 AX463227 1059 bp DNA linear PAT 15-JUL-2002
 LOCUS AX463227 Sequence 1 from Patent WO0248358.
 DEFINITION AX463227
 ACCESSION AX463227 GI:21886199
 VERSION AX463227.1 GI:21886199
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Okigami, H.
 TITLE Regulation of human chemokine-like receptor
 JOURNAL Patent: WO 0248358-A 1 20-JUN-2002;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 source 1..1059
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 99.6%; Score 1057.4; DB 6; Length 1059;
 Best Local Similarity 99.9%; Pred. No. 6.8e-246;
 Matches 1058; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGAGACACACGACGCCACCTCGACGCCAACAGCTCGCTGTTGGTGTCCCGGC 60
 DB 1 ATGAGACACACGACGCCACCTCGACGCCAACAGCTCGCTGTTGGTGTCCCGGC 60
 QY 61 TCGGCTTGGGCTTGGGCTTTCGTCGCGGTGTCTACTACAGCTTCTTGTGCTTCGGT 120
 DB 61 TCGGCTTGGGCTTGGGCTTTCGTCGCGGTGTCTACTACAGCTTCTTGTGCTTCGGT 120
 QY 121 TTACACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAG 180
 DB 121 TTACACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAG 180
 QY 181 TCTCTCAACAATATCTTGGACATGCTGCTGCGGACATCTTGTCTCTTTTCA 240
 DB 181 TCTCTCAACAATATCTTGGACATGCTGCTGCGGACATCTTGTCTCTTTTCA 240
 QY 241 GTGTTTGTGACTTCTGTTTGAAGATTCTTGAACATGACATGCTTCTTCTTCC 300
 DB 241 GTGTTTGTGACTTCTGTTTGAAGATTCTTGAACATGACATGCTTCTTCTTCC 300
 QY 301 GACAAGATCATAGAAGTGTGAAATCTCATTCATCCACACCTCCATATGATTA 360
 DB 301 GACAAGATCATAGAAGTGTGAAATCTCATTCATCCACACCTCCATATGATTA 360
 QY 361 CCGTTAACATGACAGGTATATGCTGTCTGACCCGCTCAAGTACACACGCTTCA 420
 DB 361 CCGTTAACATGACAGGTATATGCTGTCTGACCCGCTCAAGTACACACGCTTCA 420

QY 421 TACCCAGCCCGACCCGGAAGCATTTGTAAGTTTATATCATCTGTTCTGACGAC 480
 DB 421 TACCCAGCCCGACCCGGAAGCATTTGTAAGTTTATATCATCTGTTCTGACGAC 480
 QY 481 ATCCCTATTAATCTGTGAGCCCAACATCTGACATGAGACTATACATCAGACCTTGTGAT 540
 DB 481 ATCCCTATTAATCTGTGAGCCCAACATCTGACATGAGACTATACATCAGACCTTGTGAT 540
 QY 541 CAGTCTCATCTGATTCATCCATGCTTCAACCGCTTACTGTGTGCTTCTTCTTCC 600
 DB 541 CAGTCTCATCTGATTCATCCATGCTTCAACCGCTTACTGTGTGCTTCTTCTTCC 600
 QY 601 ATCTTAAGTCAATCATTTGTAAGAGCTCAGAGGAGAGCAATTTTGTCTCCGTGAC 660
 DB 601 ATCTTAAGTCAATCATTTGTAAGAGCTCAGAGGAGAGCAATTTTGTCTCCGTGAC 660
 QY 661 TACTCCAGGGGGAAGACACCGCATCTTGTTCACCATTAATCTTCTTGTGACACTT 720
 DB 661 TACTCCAGGGGGAAGACACCGCATCTTGTTCACCATTAATCTTGTGACACTT 720
 QY 721 TGGGCCCCCGCATCATCATGATTTCTTACACCTCTATGAGGCGCCCATCCGAACCCG 780
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTACACCTCTATGAGGCGCCCATCCGAACCCG 780
 QY 781 TGGCTGTGACATCATGTCGACATGTCGACATGTCGACATGTCGACATGTCGACAT 840
 DB 781 TGGCTGTGACATCATGTCGACATGTCGACATGTCGACATGTCGACATGTCGACAT 840
 QY 841 AACTTCTTCTCTACTGCTTCTTCACTGACGACAGCGGTTCCGACCATGACCGGACGCTC 900
 DB 841 AACTTCTTCTCTACTGCTTCTTCACTGACGACAGCGGTTCCGACCATGACCGGACGCTC 900
 QY 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTTACCAATCATTAATTCTTCC 960
 DB 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTTACCAATCATTAATTCTTCC 960
 QY 961 ATAAACAAGTACCCCTGGATCTTCCGCGCAACTCACTGATCAAGATGCTGTGTAC 1020
 DB 961 ATAAACAAGTACCCCTGGATCTTCCGCGCAACTCACTGATCAAGATGCTGTGTAC 1020
 QY 1021 CAGTATGACAAAATGAGAAAACCTATTAAGTATCCCG 1059
 DB 1021 CAGTATGACAAAATGAGAAAACCTATTAAGTATCCCG 1059
 RESULT 9
 AX664703 1062 bp DNA linear PAT 22-MAR-2003
 LOCUS AX664703 Sequence 12 from Patent WO02074960.
 DEFINITION AX664703
 ACCESSION AX664703 GI:29164463
 VERSION AX664703.1 GI:29164463
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Leiby, K.R., Kapeller-Liberman, R. and Glucksmann, M.
 TITLE 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
 JOURNAL of human proteins and uses thereof
 Patent: WO 02074960-A 12 26-SEP-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..1062
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 98.9%; Score 1050.8; DB 6; Length 1062;
 Best Local Similarity 99.3%; Pred. No. 2.7e-244;
 Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 1 ATGAGACACAGCAGCCACCTCGCAGCAACAGCTGCTGTTGTTGTTCCCGGC 60
Db 1 ATGAGACACAGCAGCCACCTCGCAGCAACAGCTGCTGTTGTTGTTCCCGGC 60
QY 61 TCGGCTGCGGCTTGGGTTTGTGTCCTGCTTACTACAGCCTCTTGTGCTGCTGCT 120
Db 61 TCGGCTGCGGCTTGGGTTTGTGTCCTGCTTACTACAGCCTCTTGTGCTGCTGCT 120
QY 121 TTACAGCAAAATCTTGAACAGATATCTCTTCCAGCTGTGGCAGAGACAGAG 180
Db 121 TTACAGCAAAATCTTGAACAGATATCTCTTCCAGCTGTGGCAGAGACAGAG 180
QY 121 TTACAGCAAAATCTTGAACAGATATCTCTTCCAGCTGTGGCAGAGACAGAG 180
Db 121 TTACAGCAAAATCTTGAACAGATATCTCTTCCAGCTGTGGCAGAGACAGAG 180
QY 181 TCTCTCAACAATCTTCTTGGCACTGCTGCTCCAGACATCTTGATCTCTTTTATA 240
Db 181 TCTCTCAACAATCTTCTTGGCACTGCTGCTCCAGACATCTTGATCTCTTTTATA 240
QY 241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGACATGCTCAGTCCC 300
Db 241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGACATGCTCAGTCCC 300
QY 301 GACAAGATCATAGAAAGTGTGAATTCATCCATCCACCTCCATATGATTAATCTGTA 360
Db 301 GACAAGATCATAGAAAGTGTGAATTCATCCATCCACCTCCATATGATTAATCTGTA 360
QY 361 CCGTTAAACATTAAGAGATATGCTGCTGCTGACCCGCTCAAGTACACACAGTCTGA 420
Db 361 CCGTTAAACATTAAGAGATATGCTGCTGCTGACCCGCTCAAGTACACACAGTCTGA 420
QY 421 TACCCAGCCGCGACCCGGAAGATCATTTGTAAGTGTTAATCATCACTGCTTCTGACAC 480
Db 421 TACCCAGCCGCGACCCGGAAGATCATTTGTAAGTGTTAATCATCACTGCTTCTGACAC 480
QY 481 ATCCCTATTAATGTTGTTGCTGCTCAATCTGGAAGTGAAGTGAAGTGAAGTGAAGT 540
Db 481 ATCCCTATTAATGTTGTTGCTGCTCAATCTGGAAGTGAAGTGAAGTGAAGTGAAGT 540
QY 541 CAGTCTCTATCTGGAATCCATCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CAGTCTCTATCTGGAATCCATCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 ATCTTGAACATCATTTGTTGTAAGATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 660
Db 601 ATCTTGAACATCATTTGTTGTAAGATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 660
QY 661 TACTCAGCGGGAAGACACCGCATCTTGTTCACCATTAATCTGCTGCTGCTGCTGCT 720
Db 661 TACTCAGCGGGAAGACACCGCATCTTGTTCACCATTAATCTGCTGCTGCTGCTGCT 720
QY 721 TGGGCCCCCGCATCATGATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 780
Db 721 TGGGCCCCCGCATCATGATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 780
QY 781 TGGCTGTGATCATGATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 840
Db 781 TGGCTGTGATCATGATCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 840
QY 841 AACTCTCTCTTACTGCTTATCAAGAGGCTTCCGACCATGAGCCGACACGCTC 900
Db 841 AACTCTCTCTTACTGCTTATCAAGAGGCTTCCGACCATGAGCCGACACGCTC 900
QY 901 AAGGCTTCTTCAATGTCAGAGACCTGATGCTTCAACCAATCAATTAATCTTTTCC 960
Db 901 AAGGCTTCTTCAATGTCAGAGACCTGATGCTTCAACCAATCAATTAATCTTTTCC 960
QY 961 ATTAACAAGTATGAGCTGATCTGCGCGCAACTCACTGATCATCAAGTCTGCTGATC 1020
Db 961 ATTAACAAGTATGAGCTGATCTGCGCGCAACTCACTGATCATCAAGTCTGCTGATC 1020
QY 1021 CAGTATGACAAATGAGAAATCTATAAAGTATCTCCGCTGA 1062
Db 1021 CAGTATGACAAATGAGAAATCTATAAAGTATCTCCGCTGA 1062
```

```
RESULT 10
AX664701 1526 bp DNA linear PAT 22-MAR-2003
LOCUS AX664701
DEFINITION Sequence 10 from Patent WO02074960.
ACCESSION AX664701
VERSION AX664701.1 GI:29164461
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Buteria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Leidy, K.R., Kapeller-Libermann, R. and Glucksmann, M.
38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof
Patent: WO 02074960-A 10 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
location/Qualifiers
1..1526
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
139..1200
/note="unamed protein product"
/codon_start=1
/protein_id="CAD80046.1"
/db_xref="GI:29164462"
/translation="MEHTHAIILANSLSLMSWSPSACGLGFPVYVYSLICICLIPAN
ILTVILISOLVARORSYNYILALADILVLPYVDFPFLBEDPILNQMOPVPEK
IIEVLEBSISHTSITWIPVILIDIRYAVCHPLKHTYSYFARKRYIVSYITCFPLTS
IPYVWPNITBEDISTSVHVLVWCFIVYVPCSEIFILNSIVYKLRKSNFEL
CGISTGKTIILFTITISFALIMAPRIIMLYHLYGAPIONRMVHIMSDIAMMLALL
NTAINEFLYCFISIRFRTMAATLKAFKCKOPVDFYTHNFSLTSSPWISPSNSHC
IKMLVYDKNKPIKIVSP"
ORIGIN
Query Match 98.9%; Score 1050.8; DB 6; Length 1526;
Best Local Similarity 99.3%; Pred. No. 2.6e-244;
Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGAGACACAGCAGCCACCTCGCAGCAACAGCTGCTGTTGTTGTTCCCGGC 60
Db 139 ATGAGACACAGCAGCCACCTCGCAGCAACAGCTGCTGTTGTTGTTCCCGGC 198
QY 61 TCGGCTGCGGCTTGGGTTTGTGTCCTGCTTACTACAGCCTCTTGTGCTGCTGCT 120
Db 199 TCGGCTGCGGCTTGGGTTTGTGTCCTGCTTACTACAGCCTCTTGTGCTGCTGCT 258
QY 121 TTACAGCAAAATCTTGAACAGATATCTCTTCCAGCTGTGGCAGAGACAGAG 180
Db 259 TTACAGCAAAATCTTGAACAGATATCTCTTCCAGCTGTGGCAGAGACAGAG 318
QY 181 TCTCTCAACAATCTTCTTGGCACTGCTGCTCCAGACATCTTGATCTCTTTTATA 240
Db 319 TCTCTCAACAATCTTCTTGGCACTGCTGCTCCAGACATCTTGATCTCTTTTATA 378
QY 241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGACATGCTCAGTCCC 300
Db 379 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGACATGCTCAGTCCC 438
QY 301 GACAAGATCATAGAAAGTGTGAATTCATCCATCCACCTCCATATGATTAATCTGTA 360
Db 439 GACAAGATCATAGAAAGTGTGAATTCATCCATCCACCTCCATATGATTAATCTGTA 498
QY 361 CCGTTAAACATTAAGAGATATGCTGCTGCTGACCCGCTCAAGTACACACAGTCTGA 420
Db 499 CCGTTAAACATTAAGAGATATGCTGCTGCTGACCCGCTCAAGTACACACAGTCTGA 558
QY 421 TACCCAGCCGCGACCCGGAAGATCATTTGTAAGTGTTAATCATCACTGCTTCTGACAC 480
Db 559 TACCCAGCCGCGACCCGGAAGATCATTTGTAAGTGTTAATCATCACTGCTTCTGACAC 618
QY 481 ATCCCTATTAATGTTGTTGCTGCTCAATCTGGAAGTGAAGTGAAGTGAAGTGAAGT 540
```

Db 619 ATCCCTATTCAGTGGGCGGCAACATTTGAGCTGAGACATCACTACAGACCTCTTGAT 678
QY 541 CAGCTCTCATCTGGATGCACTGCTTACCGCTTACCTGGTGCCTGCTCATCTTCTTC 600
Db 679 CAGCTCTCATCTGGGCTCACTGCTTACCGCTTACCGAGGTGCCCTGCTCATCTTCTTC 738
QY 601 ATCTTGAATCATCTATGTTGTACAAAGCTCAGAGAGAGAGCAATTTTGTCTTCGGTGC 660
Db 739 ATCTTGAATCATCTATGTTGTACAAAGCTCAGAGAGAGCAATTTTGTCTTCGGTGC 798
QY 661 TACTCAGGGGAGAGACCAACCGCATCTTGTTCACCTTACCTTCACTTCTTGCACACT 720
Db 799 TACTCAGGGGAGAGACCAACCGCATCTTGTTCACCTTACCTTCACTTCTTGCACACT 858
QY 721 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTATGAGGCGCCCATCAGAAACCG 780
Db 859 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTATGAGGCGCCCATCAGAAACCG 918
QY 781 TGGCTGTGACATCATGTTCCGACATTTGCCAATGCTAGCTTCTTGAAACAGCCATC 840
Db 919 TGGCTGTGACATCATGTTCCGACATTTGCCAATGCTAGCTTCTTGAAACAGCCATC 978
QY 841 AACTTCTCTCTATGCTTCTCATCAGCAGCGGTTCCGACATGAGCGCGCAGCGCTC 900
Db 979 AACTTCTCTCTATGCTTCTCATCAGCAGCGGTTCCGACATGAGCGCGCAGCGCTC 1038
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACCACTCAATCTTCTTC 960
Db 1039 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACCACTCAATCTTCTTC 1098
QY 961 ATTAACAAGTACCGCTTGCATCTGCGCGGCAACTCAGCTGCTGCTGCTGCTGCTAC 1020
Db 1099 ATTAACAAGTACCGCTTGCATCTGCGCGGCAACTCAGCTGCTGCTGCTGCTGCTAC 1158
QY 1021 CAGTATGACAAATAAGTAAAGTATCCCGGTGA 1062
Db 1159 CAGTATGACAAATAAGTAAAGTATCCCGGTGA 1200

RESULT 11
AX254975 1119 bp DNA 1linear PAT 10-OCT-2001
LOCUS
DEFINITION
AX254975 Sequence 19 from Patent W00170978.
VERSION
AX254975.1 GI:16074503
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
1 Taupier, R.J., Majumder, K., Spaderina, S.K., Smithson, G., Mezes, P.S.
and Verret, C.A.
TITLE
Polypeptides and nucleic acids encoding same
JOURNAL
Patent: WO 0170978-A 19 27-SEP-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..1119
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.7%; Score 1048.4; DB 6; Length 1119;
Best Local Similarity 99.9%; Pred. No. 1e-243;
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGCCCACTTGCAGCCCAACAGCTCGCTGTTGGTGTGCTCCCGCGC 60
Db 1 ATGAGACACACGACGCCCACTTGCAGCCCAACAGCTCGCTGTTGGTGTGCTCCCGCGC 60
QY 61 TCGGCTGCGGGCTTGGGTTTGTGCGCGGTGTCTACTACAGCTCTTGTGCTGCTCGGT 120

Db 61 TCGGCTGCGGGCTTGGGTTTGTGCGCGGTGTCTACTACAGCTCTTGTGCTGCTCGGT 120
QY 121 TTAACGACAAATATCTTGAACAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAG 180
Db 121 TTAACGACAAATATCTTGAACAGTATCATCTCTCCAGCTGGTGGCAAGAGAGAG 180
QY 181 TCTCTTACACATATCTTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 181 TCTCTTACACATATCTTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 GGTGTTGGGACTCTCTGTTGGAGATTTCATCTTGAATGAGATGAGTGCAGGCTCCG 300
Db 241 GGTGTTGGGACTCTCTGTTGGAGATTTCATCTTGAATGAGATGAGTGCAGGCTCCG 300
QY 301 GACAAATCATAGAAAGTGTGAAATTCATCATCATCAGACCTCATATGATTAATGTA 360
Db 301 GACAAATCATAGAAAGTGTGAAATTCATCATCATCAGACCTCATATGATTAATGTA 360
QY 361 CCGTTAACATTTGACAGGTAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 CCGTTAACATTTGACAGGTAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 TACCCAGCCCGACCCGGAAGTCAATTTGATTAATGATTAATGATTAATGATTAATG 480
Db 421 TACCCAGCCCGACCCGGAAGTCAATTTGATTAATGATTAATGATTAATGATTAATG 480
QY 481 ATCCCTATTAATGATGAGTGCACCAATCTGGAATGATTAATGATTAATGATTAATG 540
Db 481 ATCCCTATTAATGATGAGTGCACCAATCTGGAATGATTAATGATTAATGATTAATG 540
QY 541 CAGCTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 CAGCTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 601 ATCTTGAATCATCTATGTTGTACAAAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGC 660
Db 601 ATCTTGAATCATCTATGTTGTACAAAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGC 660
QY 661 TACTCAGGGGAGAGACCAACCGCATCTTGTTCACCATTAACCTCTTGTGCACTT 720
Db 661 TACTCAGGGGAGAGACCAACCGCATCTTGTTCACCATTAACCTCTTGTGCACTT 720
QY 721 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTTATGAGGCGCCATCCAGAACCG 780
Db 721 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTTATGAGGCGCCATCCAGAACCG 780
QY 781 TGGCTGTGACATCATGTTCCGACATTTGCCAATGCTAGCTTCTTGAAACAGCCATC 840
Db 781 TGGCTGTGACATCATGTTCCGACATTTGCCAATGCTAGCTTCTTGAAACAGCCATC 840
QY 841 AACTTCTCTCTATGCTTCTCATCAGCAGCGGTTCCGACATGAGCGCGCAGCGCTC 900
Db 841 AACTTCTCTCTATGCTTCTCATCAGCAGCGGTTCCGACATGAGCGCGCAGCGCTC 900
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACCACTCAATCTTCTTC 960
Db 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACCACTCAATCTTCTTC 960
QY 961 ATTAACAAGTACCGCTTGCATCTGCGCGGCAACTCAGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 ATTAACAAGTACCGCTTGCATCTGCGCGGCAACTCAGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CAGTATGACAAATAAGTAAAGTATCCCGGTGA 1050
Db 1021 CAGTATGACAAATAAGTAAAGTATCCCGGTGA 1050

RESULT 12
AX375235 1130 bp DNA 1linear PAT 01-MAR-2002
LOCUS
DEFINITION
AX375235 Sequence 25 from Patent W00210387.


```

VERSION      AX375235.1  GI:19169953
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1
Thornton, M., Paterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
Elliot, V.S., Rankummar, D., Baughn, M.R., Kallick, D.A., Walla, N.K.,
Hafalla, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L.,
Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.,
G-Protein coupled receptors
Patent: WO 0210387-A 25 07-FEB-2002;
JOURNAL
Incyte Genomics, Inc. (US)
FEATURES
source      1..1130
             Location/Qualifiers
             1..1130
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"
             /note="Incyte ID No: 7474767CBI"
ORIGIN
Query Match      98.7%; Score 1048.4; DB 6; Length 1130;
Best Local Similarity 99.9%; Pred. No. 1e-243;
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGGACACGACGCGCCACCTCGACCAAGCTCGCTGTTGGTGTCCCGCGC 60
DB 12 ATGAGGACACGACGCGCCACCTCGACCAAGCTCGCTGTTGGTGTCCCGCGC 71
QY 61 TCGGCGCTGCGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTTGTGCTCGCTCGGT 120
DB 72 TCGGCGCTGCGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTTGTGCTCGCTCGGT 131
QY 121 TTACGACAAATATCTTGACAGTATATCTCTCCAGTGTGGTGGCAAGACGAG 180
DB 132 TTACGACAAATATCTTGACAGTATATCTCTCCAGTGTGGTGGCAAGACGAG 191
QY 181 TCTCTCTACACATCTCTTGGGACCTGCTGCTGCGGACATCTTGCTCTTTTTCATA 240
DB 192 TCTCTCTACACATCTCTTGGGACCTGCTGCTGCGGACATCTTGCTCTTTTTCATA 251
QY 241 GTGTTTGTGACCTTCCGTTTGGAGATTTCATTTTAAATGATGACATGCTTCCGCT 300
DB 252 GTGTTTGTGACCTTCCGTTTGGAGATTTCATTTTAAATGATGACATGCTTCCGCT 311
QY 301 GACAAATATAGAAAGTGTGAAATCTCATCATCACAACCTCCATATGATTACTGTA 360
DB 312 GACAAATATAGAAAGTGTGAAATCTCATCATCACAACCTCCATATGATTACTGTA 371
QY 361 CCGTTAAACATTTGACAGGATATCGGTCTGCTGCGACCGGCTCAAGTACACAGGCTCA 420
DB 372 CCGTTAAACATTTGACAGGATATCGGTCTGCTGCGACCGGCTCAAGTACACAGGCTCA 431
QY 421 TTACCCAGCCCGACCCGGAAGTCAATTGTAAGTTTTCATCACTGCTTCTGACGAC 480
DB 432 TTACCCAGCCCGACCCGGAAGTCAATTGTAAGTTTTCATCACTGCTTCTGACGAC 491
QY 481 ATCCCTTATTTACTGTTGGGCCCAATCTGGAAGTGAAGTCAATCAGACACCTGTGAT 540
DB 492 ATCCCTTATTTACTGTTGGGCCCAATCTGGAAGTGAAGTCAATCAGACACCTGTGAT 551
QY 541 CACGTCCTATCTGATCCACTGCTTCAACCGTCTACCTGCTGCTGCTGCTGCTGCTGCT 600
DB 552 CACGTCCTATCTGATCCACTGCTTCAACCGTCTACCTGCTGCTGCTGCTGCTGCTGCT 611
QY 601 ATCTTGAATCAATATTTGTGTAAGCTCAGAGGAAGACAAATTTTCTGCTCGTGGC 660
DB 612 ATCTTGAATCAATATTTGTGTAAGCTCAGAGGAAGACAAATTTTCTGCTCGTGGC 671
QY 661 TACTCAGAGGGAAGACACCGCATCTTTGTCACATTAACCTCATCTTTGACACTT 720
DB 672 TACTCAGAGGGAAGACACCGCATCTTTGTCACATTAACCTCATCTTTGACACTT 731

```

```

QY 721 TGGCCCCCCCCGATCATATGATTTCTTACCACTCTATGAGGCGCCCATCCAGAACGCG 780
DB 732 TGGCCCCCCCCGATCATATGATTTCTTACCACTCTATGAGGCGCCCATCCAGAACGCG 791
QY 781 TGGCTGTGACATCATATGTCGACATTTGGCAATGTCAGTACGCTTCTGTAACAGCCATC 840
DB 792 TGGCTGTGACATCATATGTCGACATTTGGCAATGTCAGTACGCTTCTGTAACAGCCATC 851
QY 841 AACTTTTCTCTTACTGCTTCTATCAGCAAGCGGTTCCGACCATGAGCGCCGACGCTC 900
DB 852 AACTTTTCTCTTACTGCTTCTATCAGCAAGCGGTTCCGACCATGAGCGCCGACGCTC 911
QY 901 AAGCTTTCTTCAAGGCGACAGAGAACCTTGACAGTCTTACCAATCATTAATCTTTTCC 960
DB 912 AAGCTTTCTTCAAGTCTTCAAGGCGACAGAGAACCTTGACAGTCTTACCAATCATTAATCTTTTCC 971
QY 961 ATTAACAGTAGCCCTGTGATCTTCCGCGCAAACTCACAAGTCAATCAAGATGCTGATAC 1020
DB 972 ATTAACAGTAGCCCTGTGATCTTCCGCGCAAACTCACAAGTCAATCAAGATGCTGATAC 1031
QY 1021 CAGTATGACAAAATGAAAACCTTATAAA 1050
DB 1032 CAGTATGACAAAATGAAAACCTTATAAA 1061

RESULT 13
AX463235
LOCUS      AX463235
DEFINITION Sequence 9 from Patent WO0248358.
ACCESSION AX463235
VERSION    AX463235.1 GI:21886203
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1
Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Okigami, H.,
Regulation of human chemokine-like receptor
Patent: WO 0248358-A 9 20-JUN-2002;
JOURNAL     Bayer Aktiengesellschaft (DE)
FEATURES
source     1..1062
            Location/Qualifiers
            1..1062
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match      89.0%; Score 945.2; DB 6; Length 1062;
Best Local Similarity 93.1%; Pred. No. 1.2e-218;
Matches 989; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1 ATGAGGACACGACGCGCCACCTCGACCAAGCTCGCTGTTGGTGTCCCGCGC 60
DB 1 ATGAGGACACGACGCGCCACCTCGACCAAGCTCGCTGTTGGTGTCCCGCGC 60
QY 61 TCGGCGCTGCGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTTGTGCTCGCTCGGT 120
DB 61 TCGGCGCTGCGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTTGTGCTCGCTCGGT 120
QY 121 TTACGACAAATATCTTGACAGTATATCTCTCCAGTGTGGTGGCAAGACAGAG 180
DB 121 TTACGACAAATATCTTGACAGTATATCTCTCCAGTGTGGTGGCAAGACAGAG 180
QY 181 TCTCTCTAACAATATCTTGGGACCTGCTGCTGCGGACATCTTGTGCTCTTTTTCATA 240
DB 181 TCTCTCTAACAATATCTTGGGACCTGCTGCTGCGGACATCTTGTGCTCTTTTTCATA 240
QY 241 GTGTTTGTGACCTTCCGTTTGGAGATTTCATTTTAAATGATGACATGCTTCCGCT 300
DB 241 GTGTTTGTGACCTTCCGTTTGGAGATTTCATTTTAAATGATGACATGCTTCCGCT 300

```


QY 301 GACAGATCATAGAGTGTGGAATTCATCATCATCAGACCTCCATATGATTACTGTA 360
 |||
 Db 301 GACAGATCATAGAGTGTGGAATTCATCATCATCAGACCTCCATATGATTACTGTA 360
 QY 361 CCGTTAACCATGACAGATATCGCTGTGCGACCCGCTCAAGTACACACAGGTCTCA 420
 |||
 Db 361 CCGTTAACCATGACAGATATCGCTGTGCGACCCGCTCAAGTACACACAGGTCTCA 420
 QY 421 TACCCAGCCCGACCCGGAAGTCAATGTTAAGTGTAACTCACTGCTCTGACACG 480
 |||
 Db 421 TACCCAGCCCGACCCGGAAGTCAATGTTAAGTGTAACTCACTGCTCTGACACG 480
 QY 481 ATCCCTATTAATCGGTGGCCCAATCTGACATGAAAGTCACTACAGCACTGTGCAAT 540
 |||
 Db 481 ATCCCTATTAATCGGTGGCCCAATCTGACATGAAAGTCACTACAGCACTGTGCAAT 540
 QY 541 CAGCTCTCATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 600
 |||
 Db 541 CAGCTCTCATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 600
 QY 601 ATCTTGAATCAATCAATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 660
 |||
 Db 601 ATCTTGAATCAATCAATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 660
 QY 661 TACTCAGCGGGAAGACACCGGCACTGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAG 720
 |||
 Db 661 TACTCAGCGGGAAGACACCGGCACTGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAG 720
 QY 721 TGGGCCCCCGGCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 780
 |||
 Db 721 TGGGCCCCCGGCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 780
 QY 781 TGGCTGTGCAATCATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 840
 |||
 Db 781 TGGCTGTGCAATCATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 840
 QY 841 AACTTCTCTCATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 900
 |||
 Db 841 AACTTCTCTCATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 900
 QY 901 AAGGCTTTCTTCAAGTGCAGAGCAAGCACTGATCTGATCTGATCTGATCTGATCTGAT 960
 |||
 Db 901 AAGGCTTTCTTCAAGTGCAGAGCAAGCACTGATCTGATCTGATCTGATCTGATCTGAT 960
 QY 961 ATPAACAAGTAGACCCCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1020
 |||
 Db 961 ATPAACAAGTAGACCCCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1020
 QY 1021 CAGTATGACAAAATGAAAACTTATAAAGTATCCCGGTGA 1062
 |||
 Db 1021 CAGTATGACAAAATGAAAACTTATAAAGTATCCCGGTGA 1062
 RESULT 14
 CQ737667 957 bp DNA linear PAT 03-FEB-2004
 LOCUS CQ737667 Sequence 23601 from Patent WO02068579.
 DEFINITION CQ737667
 ACCESSION CQ737667
 VERSION CQ737667.1 GI:42336565
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kite, such as nucleic acid arrays, comprising a majority of
 humenons or transcripts, for detecting expression and other uses
 thereof
 Patent: WO 02068579-A 23601 06-SEP-2002;
 PE Corporation (NY) (US)
 Location/Qualifiers
 1. 957
 source

ORIGIN

/organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

Query Match 88.5%; Score 940.2; DB 6; Length 957;
 Best Local Similarity 99.7%; Pred. No. 2e-217;
 Matches 942; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 118 GGTTCACGACAAATATCTTGAACAGTATCATCTCTCCACCTGGTGCAAGAACAG 177
 |||
 Db 13 GGTTCACGACAAATATCTTGAACAGTATCATCTCTCCACCTGGTGCAAGAACAG 72
 QY 178 AAGTCTCTTCAACATATCTTGGCACTGCTGCTGCCGACATCTTGCTCTTTTC 237
 |||
 Db 73 AAGTCTCTTCAACATATCTTGGCACTGCTGCTGCCGACATCTTGCTCTTTTC 132
 QY 238 ATAGTGTGTGACCTCTGTTGGAAGTTCATCTTGAACATGACAGATGCTCAGATC 297
 |||
 Db 133 ATAGTGTGTGACCTCTGTTGGAAGTTCATCTTGAACATGACAGATGCTCAGATC 192
 QY 298 CCGACAAAGATCATAGAGTGTGGAATTTCTCATCTCAGACACCTCATATGATTACT 357
 |||
 Db 193 CCGACAAAGATCATAGAGTGTGGAATTTCTCATCTCAGACACACCTCATATGATTACT 252
 QY 358 GTACCGTTAACATGACAGATATTCGCTGTGCGACCCGCTCAAGTACACACGCTC 417
 |||
 Db 253 GTACCGTTAACATGACAGATATTCGCTGTGCGACCCGCTCAAGTACACACGCTC 312
 QY 418 TCATACCCAGCCGGAACCGGGAAGTCAATGTAAGTGTAAATGATCAGCTGCTCCGACC 477
 |||
 Db 313 TCATACCCAGCCGGAACCGGGAAGTCAATGTAAGTGTAAATGATCAGCTGCTCCGACC 372
 QY 478 AGCATCCCCCTTATCTGATGAGCCCAACATCTGAGTGAAGTCAATGATCAGACCTCTG 537
 |||
 Db 373 AGCATCCCCCTTATCTGATGAGCCCAACATCTGAGTGAAGTCAATGATCAGACCTCTG 432
 QY 538 CATCAGCTCTTATCTGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 597
 |||
 Db 433 CATCAGCTCTTATCTGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 492
 QY 598 TTCACTTGAATCAATCATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 657
 |||
 Db 493 TTCACTTGAATCAATCATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 552
 QY 658 GGTCTCTCAGCGGGAAGACACCGGCACTGTTTCAACATTAACCTCCATCTTTGCCACA 717
 |||
 Db 553 GGTCTCTCAGCGGGAAGACACCGGCACTGTTTCAACATTAACCTCCATCTTTGCCACA 612
 QY 718 CTTTGGGCCCCCGGCACTCATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 777
 |||
 Db 613 CTTTGGGCCCCCGGCACTCATGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 672
 QY 778 CGCTGCTGTGCAATCATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 837
 |||
 Db 673 CGCTGCTGTGCAATCATGTTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 732
 QY 838 ATCAACTTCTTCTTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 897
 |||
 Db 733 ATCAACTTCTTCTTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 792
 QY 898 CTCAGAGCTTTCTTCAAGTGCAGAGCAAGCACTGTTAAGTGTAACTCAATCATTAATCT 957
 |||
 Db 793 CTCAGAGCTTTCTTCAAGTGCAGAGCAAGCACTGTTAAGTGTAACTCAATCATTAATCT 852
 QY 958 TCCATTAACAAGTAGACCCCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1017
 |||
 Db 853 TCCATTAACAAGTAGACCCCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 912
 QY 1018 TACAGTATGACAAAATGAAAACTTATAAAGTATCCCGGTGA 1062
 |||
 Db 913 TACAGTATGACAAAATGAAAACTTATAAAGTATCCCGGTGA 957

RESULT 15
AL139235 0/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL139235 Accession AL139235
Fragment Name Begin End
AL139235_0 1 110000
AL139235_1 100001 210000
AL139235_2 200001 310000
AL139235_3 300001 407025
LOCUS AL139235 407025 bp DNA linear HTG 16-JUL-2001
DEFINITION Homo sapiens chromosome 9 clone RP11-3215, 31 unordered pieces.
ACCESSION AL139235
VERSION AL139235.22 GI:11034497
KEYWORDS HTG; HTGS_Phasel; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1
REFERENCE
AUTHORS Plumb,B.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
request: clonerequest@sanger.ac.uk
On Oct 26, 2000 this sequence version replaced gi:11024977.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: ba3215
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 389894 bases at least Q40
Consensus quality: 395933 bases at least Q30
Consensus quality: 400050 bases at least Q20
Insert size: 404025; sum-of-contigs
Insert size: 174695; 17.7% error; agarose-fp
Quality coverage: 4.93x in Q20 bases; sum-of-contigs Quality
coverage: 14.49x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2425: contig of 2425 bp in length
2426 2525: gap of 100 bp
2526 10023: contig of 7498 bp in length
10024 10124: gap of 100 bp
10124 36364: contig of 26241 bp in length
36365 36464: gap of 100 bp
36465 43873: contig of 7499 bp in length
43874 43973: gap of 100 bp
43974 49988: contig of 6015 bp in length
49989 50088: gap of 100 bp
50089 93603: contig of 43515 bp in length
93604 93703: gap of 100 bp
93704 96191: contig of 2488 bp in length
96192 96291: gap of 100 bp
96292 99433: contig of 3142 bp in length
99434 99533: gap of 100 bp
99534 101556: contig of 2023 bp in length
101557 101656: gap of 100 bp
101657 184109: contig of 8253 bp in length
184110 184209: gap of 100 bp

184210 290696: contig of 106487 bp in length
290697 290796: gap of 100 bp
290797 292892: contig of 2096 bp in length
292893 292992: gap of 100 bp
292993 296018: contig of 3026 bp in length
296019 296118: gap of 100 bp
296119 322154: contig of 26036 bp in length
322155 322254: gap of 100 bp
322255 324646: contig of 2392 bp in length
324647 324746: gap of 100 bp
324747 329069: contig of 4323 bp in length
329070 329169: gap of 100 bp
329170 331371: contig of 2202 bp in length
331372 331471: gap of 100 bp
331472 333909: contig of 2438 bp in length
333910 334009: gap of 100 bp
334010 369445: contig of 35436 bp in length
369446 369545: gap of 100 bp
369546 371996: contig of 2451 bp in length
371997 372096: gap of 100 bp
372097 375925: contig of 3829 bp in length
375926 376025: gap of 100 bp
376026 378132: contig of 2107 bp in length
378133 378232: gap of 100 bp
378233 381590: contig of 3358 bp in length
381591 381690: gap of 100 bp
381691 383726: contig of 2036 bp in length
383727 383826: gap of 100 bp
383827 390832: contig of 7066 bp in length
390833 390932: gap of 100 bp
390933 394216: contig of 3284 bp in length
394217 394316: gap of 100 bp
394317 396358: contig of 2042 bp in length
396359 396458: gap of 100 bp
396459 398733: contig of 2275 bp in length
398734 398833: gap of 100 bp
398834 402385: contig of 3552 bp in length
402386 402485: gap of 100 bp
402486 404522: contig of 2037 bp in length
404523 404622: gap of 100 bp
404623 407025: contig of 2403 bp in length.
Location/Qualifiers
1. .407025
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-3215"
/clone_1fb="RPC1-11.1"
1. .2425
/note="assembly fragment:05385
fragment_chain:1"
2526. .10023
/note="assembly fragment:03131
fragment_chain:1"
10124. .36364
/note="assembly fragment:06402
fragment_chain:1"
36465. .43873
/note="assembly fragment:01436
fragment_chain:1"
43974. .49988
/note="assembly fragment:04547
fragment_chain:1"
50089. .93603
/note="assembly fragment:01636
fragment_chain:1"
93704. .96191
/note="assembly fragment:02957
fragment_chain:2"
96292. .99433
/note="assembly fragment:00499
fragment_chain:2"

```

misc_feature      99534..101556
                   /note="assembly_fragment:03112
                   fragment_chain:2"
misc_feature      101657..184109
                   /note="assembly_fragment:01419
                   fragment_chain:3"
misc_feature      184210..290696
                   /note="assembly_fragment:00925
                   fragment_chain:3"
misc_feature      290797..292892
                   /note="assembly_fragment:03065
                   fragment_chain:4"
misc_feature      292993..296018
                   /note="assembly_fragment:05057
                   fragment_chain:4"
misc_feature      296119..322154
                   /note="assembly_fragment:00269"
misc_feature      322255..324646
                   /note="assembly_fragment:00666"
misc_feature      324747..329069
                   /note="assembly_fragment:01113"
misc_feature      329170..331371
                   /note="assembly_fragment:01461"
misc_feature      331472..333909
                   /note="assembly_fragment:01516"
misc_feature      334010..369445
                   /note="assembly_fragment:01942.0"
misc_feature      369546..371996
                   /note="assembly_fragment:02194"
misc_feature      372097..375925
                   /note="assembly_fragment:02249"
misc_feature      376026..378132
                   /note="assembly_fragment:02485"
misc_feature      378233..381590
                   /note="assembly_fragment:02649"
misc_feature      381691..383726
                   /note="assembly_fragment:02869"
misc_feature      383827..390832
                   /note="assembly_fragment:03540"
misc_feature      390933..394216
                   /note="assembly_fragment:04405"
misc_feature      394317..396358
                   /note="assembly_fragment:04555"
misc_feature      396459..398733
                   /note="assembly_fragment:05768"
misc_feature      398834..402385
                   /note="assembly_fragment:06106"
misc_feature      402486..404522
                   /note="assembly_fragment:06204"
misc_feature      404623..407025
                   /note="assembly_fragment:06292"

```

ORIGIN

Query Match 88.3%; Score 938; DB 2; Length 110000;
 Best Local Similarity 100.0%; Pred. No. 1.1e-216;
 Matches 938; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      125 CAGCAATATCTTGACAGTATCATCTCCAGCTGTGTGCAAGAGACAGAGTCTT 184
DB      35190 CAGCAATATCTTGACAGTATCATCTCCAGCTGTGTGCAAGAGACAGAGTCTT 35131
QY      185 CCTACAATATCTCTTGACAGTATCATCTCCAGCTGTGTGCAAGAGTCTT 244
DB      35130 CCTACAATATCTCTTGACAGTATCATCTCCAGCTGTGTGCAAGAGTCTT 35071
QY      245 TTGTGACATCTCTTGACAGTATCATCTCCAGCTGTGTGCAAGAGTCTT 304
DB      35070 TTGTGACATCTCTTGACAGTATCATCTCCAGCTGTGTGCAAGAGTCTT 35011
QY      305 AGATCATAGAGTGTGAATTTCTCATCCATCCACACTCATATGATTAAGTACCGT 364
DB      35010 AGATCATAGAGTGTGAATTTCTCATCCATCCACACTCATATGATTAAGTACCGT 34951

```

```

QY      365 TAACCATTTGACAGTATATATGCTGTCTGCAACCCGCTCAAGTACACAGCGTCTCATACC 424
DB      34950 TAACCATTTGACAGTATATGCTGTCTGCAACCCGCTCAAGTACACAGCGTCTCATACC 34891
QY      425 CAGCCCGCACCCGGAAGTATGTAAGTATTACATCACTCTTCTGACCAAGATTC 484
DB      34890 CAGCCCGCACCCGGAAGTATGTAAGTATTACATCACTCTTCTGACCAAGATTC 34831
QY      485 CCTATTATCTGTGTGCGCCCAATCTGTGACATCAATCAATCAAGCACTGTGTGACATCAG 544
DB      34830 CCTATTATCTGTGTGCGCCCAATCTGTGACATCAATCAATCAAGCACTGTGTGACATCAG 34771
QY      545 TCTCATCTGATTCACATCTGCTTACCGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
DB      34770 TCTCATCTGATTCACATCTGCTTACCGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 34711
QY      605 TGAATCATATCTGTTGTACCAAGTCAAGAGAGAGCAATTTTGTCTGCTGCTGCTGCT 664
DB      34710 TGAATCATATCTGTTGTACCAAGTCAAGAGAGAGCAATTTTGTCTGCTGCTGCTGCTGCT 34651
QY      665 CCAAGGGGAAGACCAACCGCATCTGTTTACCATTAACCTTCATCTTTGCAACATTTGGG 724
DB      34650 CCAAGGGGAAGACCAACCGCATCTGTTTACCATTAACCTTCATCTTTGCAACATTTGGG 34591
QY      725 CCCCCCGCATTCATGATTTCTTTACCACTGTATGAGGGGCGCCATCCAGAACCGCTGGC 784
DB      34590 CCCCCCGCATTCATGATTTCTTTACCACTGTATGAGGGGCGCCATCCAGAACCGCTGGC 34531
QY      785 TGGTGACATCATATGTCGACATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 844
DB      34530 TGGTGACATCATATGTCGACATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 34471
QY      845 TCTTCTCTACTGCTTTCATGACGAACGGTTTCGACCAATGCAAGCGCCGCAAGCTTAAG 904
DB      34470 TCTTCTCTACTGCTTTCATGACGAACGGTTTCGACCAATGCAAGCGCCGCAAGCTTAAG 34411
QY      905 CTTTCTTCAAGTCCAGAGCAACCTGTACATTTTACCAATCAATCAATCAATCAATCAATCA 964
DB      34410 CTTTCTTCAAGTCCAGAGCAACCTGTACATTTTACCAATCAATCAATCAATCAATCAATCA 34351
QY      965 CAAATGACCCCTGATCTGCGCGCAAACTCACTGATCAATCAATGATGATGATGATGATGAT 1024
DB      34350 CAAATGACCCCTGATCTGCGCGCAAACTCACTGATCAATCAATGATGATGATGATGATGAT 34291
QY      1025 ATGACAAAATGAAAACTATATAAGTATCCCGTGA 1062
DB      34290 ATGACAAAATGAAAACTATATAAGTATCCCGTGA 34253

```

Search completed: December 15, 2004, 00:42:52
 Job time : 3579.97 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: December 14, 2004, 21:00:04 ; Search time 497.373 Seconds
(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225B-15

Perfect score: 1062
Sequence: 1 atggagcacaacgacgccc.....ctataaagatccccgtga 1062

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_23sep04:*
1: geneseqn180a:*
2: geneseqn190a:*
3: geneseqn200a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1062	100.0	1062	6 ABK15563	Abk15563 cDNA enco
2	1062	100.0	1062	6 ABT04873	Abt04873 Human G P
3	1062	100.0	1062	10 ACC44115	Acc44115 Human AXO
4	1060.4	99.8	1062	6 ABK15562	Abk15562 cDNA enco
5	1060.4	99.8	1062	10 ACC71785	Acc71785 Human G P
6	1060.4	99.8	1202	8 ABV73364	Abv73364 Human TGR
7	1060	99.8	1776	10 ADF70587	Adf70587 Orphan re
8	1058.8	99.7	1343	4 AAS15731	Aas15731 DNA enco
9	1058.8	99.7	1343	10 ADJ87768	Adj87768 G-coupled
10	1058.8	99.7	1343	12 ADI79324	Adi79324 NOV11 cod
11	1058.8	99.7	1343	12 ADO56003	Ado56003 DNA enco
12	1058.8	99.7	2273	12 ADO28955	Ado28955 Human nov
13	1057.4	99.6	1059	6 ABN84269	Abn84269 Human che
14	1050.8	98.9	1526	6 AAC46858	Aac46858 Human 7TM
15	1050.8	98.9	1526	8 ACA60889	Aca60889 Human cDN
16	1050.8	98.9	1526	10 ABB57024	Abb57024 cDNA enco
17	1048.4	98.7	1119	4 AAS15730	Aas15730 DNA enco
18	1048.4	98.7	1119	10 ADJ87766	Adj87766 G-coupled
19	1048.4	98.7	1119	12 ADI79322	Adi79322 NOV10a co
20	1048.4	98.7	1119	12 ADO56001	Ado56001 DNA enco
21	1048.4	98.7	1130	6 AAD29672	Aad29672 Human G-P

22	1048.4	98.7	2198	9 ACF05275	Acf05275 Human G-P
23	1039.8	97.9	2189	10 ACF05276	Acf05276 Human G-P
24	1039.8	97.9	2189	10 ADD18120	Add18120 Human G-P
25	945.2	89.0	1062	6 ABN84273	Abn84273 Human che
26	940.2	88.5	957	12 ADI79347	Adi79347 NOV10b co
27	940.2	88.5	957	12 ADO56026	Ado56026 DNA enco
28	936.4	88.2	1002	6 AAD27501	Aad27501 Human G-P
29	936.4	88.2	1002	10 AAD61658	Aad61658 Human GPC
30	936.4	88.2	1158	6 AB242880	Ab242880 Human GPC
31	936	88.1	1032	6 ABN84271	Abn84271 Human che
32	936	88.1	1070	6 ABN84270	Abn84270 Human che
33	936	88.1	1826	6 ABN84272	Abn84272 Human che
34	934.8	88.0	1000	6 AAD28102	Aad28102 Human thy
35	917.2	86.4	946	10 ADJ87879	Adj87879 G-coupled
36	844	79.5	1038	8 ABV73373	Abv73373 Mouse TGR
37	844	79.5	1038	12 ADO28957	Ado28957 Mouse nov
38	842.6	79.3	864	4 AAS42811	Aas42811 Human G P
39	828.4	78.0	867	6 AAS98134	Aas98134 Human DNA
40	828.4	78.0	894	6 AAS98143	Aas98143 Human DNA
41	809.6	76.2	963	10 ADC12695	Adc12695 Human GPC
42	758	71.4	795	6 AAS98071	Aas98071 Human DNA
43	758	71.4	795	8 AB242552	Ab242552 Human G P
44	524	49.3	930	4 AAS42808	Aas42808 Human G P
45	268.2	25.3	1023	12 ADO28937	Ado28937 Mouse nov

ALIGNMENTS

RESULT 1	ABK15563
ID	ABK15563 standard; cDNA; 1062 BP.
AC	ABK15563;
XX	
DT	08-MAY-2002 (first entry)
XX	
DB	cDNA encoding novel G-protein coupled receptor TGR8, version #2.
KW	G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
KW	cytotoxic; antiinflammatory; anticancer; fetal brain;
KW	central nervous system disease; circulatory organ disorder; cancer;
KW	metabolic disease; immunological disease; gastrointestinal disease;
KW	gene therapy; transgenic animal; human; gene; ss.
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1062
FT	/*tag= a
FT	/product= "TGR8"
FT	/note= "G-protein coupled receptor"
XX	
MO	WO200194582-A1.
XX	
PD	13-DEC-2001.
XX	
PF	01-JUN-2001; 2001WO-0P004643.
PR	02-JUN-2000; 2000JP-00170446.
PR	23-JUN-2000; 2000JP-00194926.
XX	
PA	(TAKE) TAKEDA CHEM IND LTD.
XX	
PI	Terao Y, Matsui H, Shintani Y;
DR	WPI: 2002-164317/21.
XX	
PT	P-PSDB; AAU76416.
XX	
PT	Human fetal brain-originated G protein-coupled receptor protein TGR8 and
PT	encoding DNA, for developing drugs to treat e.g. diseases of the central
PT	nervous system or circulatory organs, cancer, and metabolic diseases.
XX	

PS Claim 5; Page 94-95; 102bp; Japanese.
XX
CC The invention describes a human fetal brain-originated G protein-coupled
CC receptor protein, or its salt. The protein and encoded DNA are useful for
CC developing drugs to treat e.g. diseases of the central nervous system or
CC circulatory organs, cancer, metabolic diseases, immunological diseases
CC and gastrointestinal diseases. The invention also describes creation of a
CC probe for gene therapy and construction of a transgenic animal. This
CC sequence encodes the novel G-protein coupled receptor TGR8, described in
CC the method of the invention
XX
SQ Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;
Query Match 100.0%; Score 1062; DB 6; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.4e-284;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACACACGACCGCCACCTCGACGCCAACAGCTCGCTGTTGTGTGTCCTCCCGGC 60
DB 1 ATGAGACACACGACCGCCACCTCGACGCCAACAGCTCGCTGTTGTGTGTCCTCCCGGC 60
QY 61 TCGGCTGGGGCTTGGGTTTCCGTGCGGTGTTACTAAGCTCTTGCTGTGCTCGGT 120
DB 61 TCGGCTGGGGCTTGGGTTTCCGTGCGGTGTTACTAAGCTCTTGCTGTGCTCGGT 120
QY 121 TTACGAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAGAG 180
DB 121 TTACGAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAGAG 180
QY 181 TCTCTCTACACTATCTCTGTGGACCTGCTGCTGCCAGACATCTGTGCTCTTTTCATA 240
DB 181 TCTCTCTACACTATCTCTGTGGACCTGCTGCTGCCAGACATCTGTGCTCTTTTCATA 240
QY 241 GTGTTGTGAGCTTCTCTGTGGAAGTTTCTTGAAGATGCAATGCTTCAAGTCCCTCC 300
DB 241 GTGTTGTGAGCTTCTCTGTGGAAGTTTCTTGAAGATGCAATGCTTCAAGTCCCTCC 300
QY 301 GACAAGATCATGAAAGTGTGGAATCTCATCATCAACACCTCCATATGATTAATCTGTA 360
DB 301 GACAAGATCATGAAAGTGTGGAATCTCATCATCAACACCTCCATATGATTAATCTGTA 360
QY 361 CCGTTTAACTTGAACAGGTATATCGCTGTGTGCCACCGCTCAAGTACACAGGTCTCA 420
DB 361 CCGTTTAACTTGAACAGGTATATCGCTGTGTGCCACCGCTCAAGTACACAGGTCTCA 420
QY 421 TACCCAGCCCGACCCGGAAAGTCAATGTAAGTGTTAATACCTGCTCTTCTGACACAG 480
DB 421 TACCCAGCCCGACCCGGAAAGTCAATGTAAGTGTTAATACCTGCTCTTCTGACACAG 480
QY 481 ATCCCTTATTAATGTTGTCGCCAACAATCTGACCTGAAGACTATCATGACACCTCTGAT 540
DB 481 ATCCCTTATTAATGTTGTCGCCAACAATCTGACCTGAAGACTATCATGACACCTCTGAT 540
QY 541 CAGGTCCTCATGTGATCCACTGCTTACCGTCTAAGTGTGCTCTGCTCATCTTCTTC 600
DB 541 CAGGTCCTCATGTGATCCACTGCTTACCGTCTAAGTGTGCTCTGCTCATCTTCTTC 600
QY 601 ACTTGAATCAATCATGTTGTAAAGCTCAGAGAGAAAGAAATTTTCGTCTCCGTGGC 660
DB 601 ACTTGAATCAATCATGTTGTAAAGCTCAGAGAGAAAGAAATTTTCGTCTCCGTGGC 660
QY 661 TACTCCAGGGGAGAGACACCGCCATCTTGTACCATTAATCTTCATCTTGGCACACTT 720
DB 661 TACTCCAGGGGAGAGACACCGCCATCTTGTACCATTAATCTTCATCTTGGCACACTT 720
QY 721 TGGGCCCCCGGATCATGATTTCTTTTACACTCTCTATGAGGCGCCCATTCAGAACGCG 780
DB 721 TGGGCCCCCGGATCATGATTTCTTTTACACTCTCTATGAGGCGCCCATTCAGAACGCG 780
QY 781 TGGGTGATGATCATGATGTCGACATTTGGCAACATGCTTAAGCTTCTGTAACACAGCATC 840
DB 781 TGGGTGATGATCATGATGTCGACATTTGGCAACATGCTTAAGCTTCTGTAACACAGCATC 840

QY 841 AACTTCTTCTCTACTGCTTATGACAGACGGTTCCGACACCATGACGCGCCACGCTC 900
DB 841 AACTTCTTCTCTACTGCTTATGACAGACGGTTCCGACACCATGACGCGCCACGCTC 900
QY 901 AAGGCTTCTTCAATGCGACAGAGCAACCTGTACAGTTCTACACCAATCATTAATTTTC 960
DB 901 AAGGCTTCTTCAATGCGACAGAGCAACCTGTACAGTTCTACACCAATCATTAATTTTC 960
QY 961 ATTAACAGTACGCTTGTATCTGCGCGCAAACTCACTGATCAAGATGCTGTGTAC 1020
DB 961 ATTAACAGTACGCTTGTATCTGCGCGCAAACTCACTGATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAAATGAAAACTTAATAAGATCCCGTGA 1062
DB 1021 CAGTATGACAAAATGAAAACTTAATAAGATCCCGTGA 1062
RESULT 2
ABT04873
ID ABT04873 standard; cDNA; 1062 BP.
XX
AC ABT04873;
XX
DT 11-OCT-2002 (first entry)
XX
DE Human G protein coupled receptor hRUP35 coding sequence.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200242461-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-US044386.
XX
PR 27-NOV-2000; 2000US-0253404P.
PR 12-DEC-2000; 2000US-0253366P.
PR 20-FEB-2001; 2001US-0270266P.
PR 20-FEB-2001; 2001US-0270286P.
PR 06-APR-2001; 2001US-0282032P.
PR 06-APR-2001; 2001US-0282356P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282365P.
PR 14-MAY-2001; 2001US-0230917P.
PR 31-JUL-2001; 2001US-0309208P.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX
XX WPI; 2002-566565/60.
XX P-PSDB; ABJ04075.
XX
PT Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
PS Claim 31; Page 70; 84p; English.
XX
CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR coding sequence of the invention
XX
SQ Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;
Query Match 100.0%; Score 1062; DB 6; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.4e-284;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 1062; DB 10; Length 1062;
 Best Local Similarity 100.0%; Pred. No. 1.4e-284;
 Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGACACGACGCGCCCACTCGAGCCAAAGCTGCTCTTGGTGGTCCCGGCG 60
DB 1 ATGAGACACGACGCGCCCACTCGAGCCAAAGCTGCTCTTGGTGGTCCCGGCG 60
QY 61 TCGGCTGCGGCTGGGTTTGGTCCCGGCTCTTCAAGCTCTTGGTGGTCCCGGCG 120
DB 61 TCGGCTGCGGCTGGGTTTGGTCCCGGCTCTTCAAGCTCTTGGTGGTCCCGGCG 120
QY 121 TTACGACCAATATCTTGAAGTATGATCTCTCCAGCTGGTGGCAAGACAGAG 180
DB 121 TTACGACCAATATCTTGAAGTATGATCTCTCCAGCTGGTGGCAAGACAGAG 180
QY 181 TCCCTCTCAACTATCTTGGCACTGGCTGGTGGCAAGCTGGTGGTCTTGGTGGT 240
DB 181 TCCCTCTCAACTATCTTGGCACTGGCTGGTGGCAAGCTGGTGGTCTTGGTGGT 240
QY 241 GTGTTTGGACCTTCTGTTGGAAGATTTCATCTTGAACATGAGATGCTCAGTCCC 300
DB 241 GTGTTTGGACCTTCTGTTGGAAGATTTCATCTTGAACATGAGATGCTCAGTCCC 300
QY 301 GACAAAGATCATGAAAGTCTGGAATTTCTATCTCAATCCACACTTCATATGATTA 360
DB 301 GACAAAGATCATGAAAGTCTGGAATTTCTATCTCAATCCACACTTCATATGATTA 360
QY 361 CCGTTAAACATGACAGGTATATGCTGTGCGACCGGCTCAAGTACCAAGGTCTCA 420
DB 361 CCGTTAAACATGACAGGTATATGCTGTGCGACCGGCTCAAGTACCAAGGTCTCA 420
QY 421 TACCAGCGCGACCGGAAAGTATGATGTTTATCATCTGCTCTGACCAAG 480
DB 421 TACCAGCGCGACCGGAAAGTATGATGTTTATCATCTGCTCTGACCAAG 480
QY 481 ATCCCTATTAAGTGGCCCAACATCTGACGTAAGACATCACTGCTCTGACCAAG 540
DB 481 ATCCCTATTAAGTGGCCCAACATCTGACGTAAGACATCACTGCTCTGACCAAG 540
QY 541 CAGCTCTCATCTGATTCACGCTTACCGTCTACCTGGTGGTCCCTCATCTTCTTC 600
DB 541 CAGCTCTCATCTGATTCACGCTTACCGTCTACCTGGTGGTCCCTCATCTTCTTC 600
QY 601 ATCTGAACATCAATGTTGTCAGAGCTCAGAGAAAGCAATTTGTTCCGTCG 660
DB 601 ATCTGAACATCAATGTTGTTCAAGCTCAGAGAAAGCAATTTGTTCCGTCG 660
QY 661 TACTCCAGGGGAAAGCAACCGCATCTGTTACCATTAACCTTTCGACACATT 720
DB 661 TACTCCAGGGGAAAGCAACCGCATCTGTTACCATTAACCTTTCGACACATT 720
QY 721 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTTATGAGGCGCCCATCAGAAC 780
DB 721 TGGGCCCCCGCATCATGATGATTTCTTACCACTCTTATGAGGCGCCCATCAGAAC 780
QY 781 TGGCTGTGACATCATGTCGACATGTCGACATGTCGACATGTCGACATGTCGAC 840
DB 781 TGGCTGTGACATCATGTCGACATGTCGACATGTCGACATGTCGACATGTCGAC 840
QY 841 AACTCTTCTCTACTGCTTATGACGAGGGGTTCCGACCATGAGCGCGCACGCTC 900
DB 841 AACTCTTCTCTACTGCTTATGACGAGGGGTTCCGACCATGAGCGCGCACGCTC 900
QY 901 AAGGCTTTCTTCAAGTCCAGAGCACTGTAAGTCTTACACCAATCAATTAATTTTCC 960
DB 901 AAGGCTTTCTTCAAGTCCAGAGCACTGTAAGTCTTACACCAATCAATTAATTTTCC 960
QY 961 ATAAAGTAAAGCTTGGATCTCGCGGCAAACTCACTGATCAAGATGCTGTGTAC 1020
DB 961 ATAAAGTAAAGCTTGGATCTCGCGGCAAACTCACTGATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAATAATGAAACTATATAAGTATCCCGGTGA 1062
DB 1021 CAGTATGACAAATAATGAAACTATATAAGTATCCCGGTGA 1062

```

DB 1021 CAGTATGACAAATAATGAAACTATATAAGTATCCCGGTGA 1062

RESULT 4
 ID ABK15562 standard; cDNA; 1062 BP.
 XX ABK15562;
 AC
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE cDNA encoding novel G-protein coupled receptor TGR8, version #1.
 XX
 KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
 KW cytosolic; antiinflammatory; antidiabetic; fetal brain;
 KW central nervous system disease; circulatory organ disorder; cancer;
 KW metabolic disease; immunological disease; gastrointestinal disease;
 KW gene therapy; transgenic animal; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH
 FT Key Location/Qualifiers
 FT CDS 1..1062
 FT /tag= a
 FT /product= "TGR8"
 FT /note= "G-protein coupled receptor"
 XX
 PN WO200194582-A1.
 XX
 PD 13-DEC-2001.
 XX
 PP 01-JUN-2001; 2001WO-JP004643.
 XX
 PR 02-JUN-2000; 2000JP-00170446.
 XX
 PR 23-JUN-2000; 2000JP-00194926.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Terao Y, Matsui H, Shintani Y;
 XX
 DR WPI; 2002-164317/21.
 XX
 DR P-PSDB; AAU76416.
 XX
 PT Human fetal brain-originated G protein-coupled receptor protein TGR8 and
 PT encoding DNA, for developing drugs to treat e.g. diseases of the central
 PT nervous system or circulatory organs, cancer, and metabolic diseases.
 XX
 PS Claim 5; Page 94; 102pp; Japanese.
 XX
 CC The invention describes a human fetal brain-originated G protein-coupled
 CC receptor protein, or its salt. The protein and encoded DNA are useful for
 CC developing drugs to treat e.g. diseases of the central nervous system or
 CC circulatory organs, cancer, metabolic diseases, immunological diseases
 CC and gastrointestinal diseases. The invention also describes creation of a
 CC probe for gene therapy and construction of a transgenic animal. This
 CC sequence encodes the novel G-protein coupled receptor TGR8, described in
 CC the method of the invention
 XX

Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;

Query Match 99.8%; Score 1060.4; DB 6; Length 1062;
 Best Local Similarity 99.9%; Pred. No. 3.9e-284;
 Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAGACACGACGCGCCCACTCGAGCCAAAGCTGCTCTTGGTGGTCCCGGCG 60
DB 1 ATGAGACACGACGCGCCCACTCGAGCCAAAGCTGCTCTTGGTGGTCCCGGCG 60
QY 61 TCGGCTGCGGCTGGGTTTGGTCCCGGCTCTTCAAGCTCTTGGTGGTCCCGGCG 120
DB 61 TCGGCTGCGGCTGGGTTTGGTCCCGGCTCTTCAAGCTCTTGGTGGTCCCGGCG 120

```


QY	121	TTACACGCAAAATATCTGACAGTGATCAATCCCTCCCGAGCTGGTGGACAAGAACAGAAAG	180
Db	121	TTACACGCAAAATATCTGACAGTGATCAATCCCTCCCGAGCTGGTGGACAAGAACAGAAAG	180
QY	181	TCCTCTCAACAATATCTCTTGGGCACTCGCTGCTGCAGACATCTTGATCTCTTTTTCATA	240
Db	181	TCCTCTCAACAATATCTCTTGGGCACTCGCTGCTGCAGACATCTTGATCTCTTTTTCATA	240
QY	241	GTGTTTGTGACCTTCTGTTGGAGATTTCACTTTGAACATGACAGATGCTCAGGTCCCC	300
Db	241	GTGTTTGTGACCTTCTGTTGGAGATTTCACTTTGAACATGACAGATGCTCAGGTCCCC	300
QY	301	GACAAAGATCATGAAAGTGTGGAAATTTCTATCTCATCTCACAACCTCCATATGATTACTGTA	360
Db	301	GACAAAGATCATGAAAGTGTGGAAATTTCTATCTCATCTCACAACCTCCATATGATTACTGTA	360
QY	361	CCGTTAACATTAAGAGTGTGGAAATTTCTATCTCATCTCACAACCTCCATATGATTACTGTA	420
Db	361	CCGTTAACATTAAGAGTGTGGAAATTTCTATCTCATCTCACAACCTCCATATGATTACTGTA	420
QY	421	TACCCAGCCCGCACCCCGGAAAGTCAATGTAAGTGTTTAATCATCACTGCTTCCCTGCACAGC	480
Db	421	TACCCAGCCCGCACCCCGGAAAGTCAATGTAAGTGTTTAATCATCACTGCTTCCCTGCACAGC	480
QY	481	ATCCCTTATTAAGTGTGGCCCAACAATCTGGAACTGAAGACTAATCAGACACCTCTGTGCAT	540
Db	481	ATCCCTTATTAAGTGTGGCCCAACAATCTGGAACTGAAGACTAATCAGACACCTCTGTGCAT	540
QY	541	CACGTCATCATGTGATCCACAGCTCAACCGTCTACCTGCTGCCCCGCTCCCATCTTCTTC	600
Db	541	CACGTCATCATGTGATCCACAGCTCAACCGTCTACCTGCTGCCCCGCTCCCATCTTCTTC	600
QY	601	ATCTTGAATCAATCAATGTTGTGACAAAGCTCAGAGGAAAGAGCAATTTTCTCCGTGGAC	660
Db	601	ATCTTGAATCAATCAATGTTGTGACAAAGCTCAGAGGAAAGAGCAATTTTCTCCGTGGAC	660
QY	661	TACTCCACGGGGAGAACCAACCGGCATCTTGTTCACCATTAACCTGCATCTTTGGCACACTT	720
Db	661	TACTCCACGGGGAGAACCAACCGGCATCTTGTTCACCATTAACCTGCATCTTTGGCACACTT	720
QY	721	TGGGGCCCCCGCATATCATGATGATTTCTTTAACCACTCTATGGGGGCGCCATCCAGAACCGC	780
Db	721	TGGGGCCCCCGCATATCATGATGATTTCTTTAACCACTCTATGGGGGCGCCATCCAGAACCGC	780
QY	781	TGGCTGTGTGCATCATGTGCCGACATTTGGCAACATGCTTAAGCCCTTCTGAACACAGCCATC	840
Db	781	TGGCTGTGTGCATCATGTGCCGACATTTGGCAACATGCTTAAGCCCTTCTGAACACAGCCATC	840
QY	841	AACTTCTTCTCTACTGCTTCAATCAGCAAGCGGTTCCGACCATGGACCGCCACAGCTC	900
Db	841	AACTTCTTCTCTACTGCTTCAATCAGCAAGCGGTTCCGACCATGGACCGCCACAGCTC	900
QY	901	AAAGCTTCTTCAATGATGCCAGAAAGCAACTGTGACAGTTCTACACAAATCATTAATCTTTCC	960
Db	901	AAAGCTTCTTCAATGATGCCAGAAAGCAACTGTGACAGTTCTACACAAATCATTAATCTTTCC	960
QY	961	ATAACAAGTAGAGCCCTGTGATCTCGCCGGGAAATCTACACTGCAATCAAGATGCTGGTGTAC	1020
Db	961	ATAACAAGTAGAGCCCTGTGATCTCGCCGGGAAATCTACACTGCAATCAAGATGCTGGTGTAC	1020
QY	1021	CAGTATGACAAAATGAGAAAACCTTATTAAGAATATCCCGTGA	1062
Db	1021	CAGTATGACAAAATGAGAAAACCTTATTAAGAATATCCCGTGA	1062
RESULT 5			
ACC1785			
ID	ACC1785 standard; DNA; 1062 BP.		
XX	ACC1785;		
XX	25-JUL-2003 (first entry)		
XX			

DE	Human G protein-coupled receptor DNA #SQ ID 1.
XX	
KW	Human; G protein-coupled receptor; anorectic; eating disorder; obesity;
KW	gene; da.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..1062
FT	/tag= a
FT	/product= "G protein-coupled receptor"
XX	
PN	MO2003027142-A1.
XX	
PD	03-APR-2003.
XX	
PF	19-SEP-2002; 2002MO-JP009626.
XX	
PR	21-SEP-2001; 2001JP-00288278.
XX	
PA	(YAMA) YAMANOUCHI PHARM CO LTD.
PI	Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T;
PI	Hiyama H;
XX	
DR	WPI; 2003-333291/31.
XX	
DR	P-Psdb; ABR44437.
XX	
PT	New G protein-coupled receptor and encoded gene, useful in screening
PT	preventives or remedies for eating disorders or obesity.
XX	
PS	Disclosure; Page 31-33; 48pp; Japanese.
XX	
CC	The invention relates to a novel G protein-coupled receptor. The protein
CC	and its encoded gene are useful for screening preventives or remedies for
CC	eating disorders or obesity. The current sequence represents a G protein
CC	coupled receptor encoding sequence
XX	
SQ	Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;
Query Match	99.8%; Score 1060.4; DB 10; Length 1062;
Best Local Similarity	99.9%; Pred. No. 3.9e-284;
Matches 1061; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Dn	
Oy	1 ATGGAGCAGACGACGCCACCCTGCAGCAAGATCGTGTGGTGTCGCCGCGC 60
Dn	1 ATGGAGCAGACGACGCCACCCTGCAGCAAGATCGTGTGGTGTCGCCGCGC 60
Oy	61 TCGGCTGCGGCTTGGTTGTCGTCGCCGCTGTACTAAGCTCTTCTGCTCGCTCGT 120
Dn	61 TCGGCTGCGGCTTGGTTGTCGTCGCCGCTGTACTAAGCTCTTCTGCTCGCTCGT 120
Oy	121 TTACGAGCAATATCTTGAACAGTAGATCATCTCTCCAGCTGTGTGGCAAGAAGACAAG 180
Dn	121 TTACGAGCAATATCTTGAACAGTAGATCATCTCTCCAGCTGTGTGGCAAGAAGACAAG 180
Oy	181 TTCCTCTCAACATATCTTGGCACTGCTGTCGCGACATCTTGTGCTCTTTTTCATA 240
Dn	181 TTCCTCTCAACATATCTTGGCACTGCTGTCGCGACATCTTGTGCTCTTTTTCATA 240
Oy	241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGAGATGCTCAGATCCCC 300
Dn	241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGAGATGCTCAGATCCCC 300
Oy	301 GACAAAGATCATGAAGTGTGGAATTCTCATCTCCACACCTTCATATGATTAATCTGA 360
Dn	301 GACAAAGATCATGAAGTGTGGAATTCTCATCTCCACACCTTCATATGATTAATCTGA 360
Oy	361 CGGTTAACATTAAGTATATGCTGTGTCGACACCGGCTCAAGTACCAAGAGGTCTCA 420
Dn	361 CGGTTAACATTAAGTATATGCTGTGTCGACACCGGCTCAAGTACCAAGAGGTCTCA 420
Oy	421 TACCAGCCGACCGGAAGTATGTATGATTTTATCATCACCTGCTTCTACAGCAGC 480

```

Db      |||||||
421  TACCCAGCCCGGACCGGAAAGTCAATGTAAGTTTACATCACTGCTCTGACCAAGC 480
Qy      |||||||
481  ATCCCTATTATCTGTTGGCCCAACATCTGAGATGAAAGCTATACAGACCTCTGAGAT 540
Db      |||||||
481  ATCCCTATTATCTGTTGGCCCAACATCTGAGATGAAAGCTATACAGACCTCTGAGAT 540
Qy      |||||||
541  CAGGCTCATCTGATGATCACTGCTTCAACCGTCTACCTGATGCTGCTCATCTTCTTC 600
Db      |||||||
541  CAGGCTCATCTGATGATCACTGCTTCAACCGTCTACCTGATGCTGCTCATCTTCTTC 600
Qy      |||||||
601  ATCTTGAATCAATCAATGTTGTGACAAAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGC 660
Db      |||||||
601  ATCTTGAATCAATCAATGTTGTGACAAAGCTCAGAGAGAGAGCAATTTTGTCTCCGTGC 660
Qy      |||||||
661  TACTCCAGGGGAGAGACCAACCGCATCTTGTTCACATTAACCTCCATCTTTCGACACTT 720
Db      |||||||
661  TACTCCAGGGGAGAGACCAACCGCATCTTGTTCACATTAACCTCCATCTTTCGACACTT 720
Qy      |||||||
721  TGGGCCCCCGCATCATCATGATTTCTTACCACTCTATGAGGCGCCCATCCAGAACCGC 780
Db      |||||||
721  TGGGCCCCCGCATCATCATGATTTCTTACCACTCTATGAGGCGCCCATCCAGAACCGC 780
Qy      |||||||
781  TGGCTGTGACATCATATGTCGACATTTGCCAACATGCTGCTTGTGAAACAGAGCATC 840
Db      |||||||
781  TGGCTGTGACATCATATGTCGACATTTGCCAACATGCTGCTTGTGAAACAGAGCATC 840
Qy      |||||||
841  AACTCTCTCTCTACGCTTCAATCAGAACGCGTTCGACACATGAGCGGACGACGCTC 900
Db      |||||||
841  AACTCTCTCTCTACGCTTCAATCAGAACGCGTTCGACACATGAGCGGACGACGCTC 900
Qy      |||||||
901  AAGCTTTCTTCAAGTCCAGAACCACTGTACAGTTTCAACCAATCAATCAATCTTTTC 960
Db      |||||||
901  AAGCTTTCTTCAAGTCCAGAACCACTGTACAGTTTCAACCAATCAATCAATCTTTTC 960
Qy      |||||||
961  ATTAACAAGTACCCCTGATCTGCGCGGCAACCTCACTGATCAAGATGCTGTGTAC 1020
Db      |||||||
961  ATTAACAAGTACCCCTGATCTGCGCGGCAACCTCACTGATCAAGATGCTGTGTAC 1020
Qy      |||||||
1021  CAGTATGACAAATAATGAAAACTATATAAGTATCCCGTGA 1062
Db      |||||||
1021  CAGTATGACAAATAATGAAAACTATATAAGTATCCCGTGA 1062

RESULT 6
ABV73364
ID   ABV73364 standard; DNA; 1202 BP.
XX
AC   ABV73364;
XX
XX
XX  22-JAN-2003 (first entry)
XX
XX  Human TGR20 polypeptide encoding DNA.
XX
XX  G-protein coupled receptor; GPCR; antiproliferative; antiinflammatory;
XX  antilipemic; neuroprotective; antidiabetic; antidiabetic; human;
XX  antiparkinsonian; antilipemic; TGR20; gene; ds.
XX
XX  Homo sapiens.
XX
XX  Key      Location/Qualifiers
XX  CDS      68..1129
XX           /*tag=
XX           /product= "TGR20"
XX           /note= "GPCR polypeptide"
XX
XX  MO200277001-A2.
XX
XX  PD      03-OCT-2002.
XX
XX  PF      08-MAR-2002; 2002MO-US007171.
XX
XX  PR      09-MAR-2001; 2001US-00802803.

```

```

PR      16-MAR-2001; 2001US-0276649P.
XX
XX  (TULIA-) TULARIK INC.
XX
XX  Tian H, Zhao J, Chen J, Cutler G;
XX
XX  WPI; 2003-018881/01.
XX
XX  P-PSDB; ABB82499.
XX
XX  New G-protein coupled receptor polypeptides and polynucleotides useful
XX  for identifying compounds for treating a TGR-associated disorder, e.g.
XX  psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
XX  disease, anemia.
XX
XX  Claim 6; Page 63; 87pp; English.
XX
XX  The invention relates to G-protein coupled receptor (GPCR) polypeptides
XX  and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
XX  TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
XX  useful for identifying compounds for treating a TGR-associated disorder,
XX  such as psoriasis, inflammatory bowel disease, hyperlipidemia,
XX  Parkinson's disease, Huntington's disease, anemia, immune and blood
XX  disorders, ulcerative colitis, Crohn's disease or kidney, liver,
XX  CC They are also useful for identifying cells such as kidney, liver,
XX  CC hypothalamus, colon, adipose, or spleen cells, for forensics and
XX  CC paternity determination, diagnosing diseases and examining signal
XX  CC transduction. The present sequence represents a human TGR20 polypeptide
XX  CC encoding DNA
XX
XX  Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;
XX
XX  Query Match      99.8%; Score 1060.4; DB 8; Length 1202;
XX  Best Local Similarity 99.9%; Pred. No. 4.1e-284;
XX  Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
Qy      1  ATGAGCAGCAGCAGCCCACTGCGACGCAAGCTGCTGTGTTGCTCCCGGCG 60
Db      68  ATGAGCAGCAGCAGCCCACTGCGACGCAAGCTGCTGTGTTGCTCCCGGCG 127
Qy      61  TGGGCTCGCGGCTTGGGTTTGGTGGCGGTGCTACAGCTTCTGCTGCTCGGT 120
Db      128  TGGGCTCGCGGCTTGGGTTTGGTGGCGGTGCTACAGCTTCTGCTGCTCGGT 187
Qy      121  TTACAGCAAAATATTTGACAGTATCATCTTCCAGCTGTGTGCAAGAAAGACAGAG 180
Db      188  TTACAGCAAAATATTTGACAGTATCATCTTCCAGCTGTGTGCAAGAAAGACAGAG 247
Qy      181  TCTCTTACAACTATCTCTTGGCACTGCTGCTGCGGCACTTGTGCTCTTTTCA 240
Db      248  TCTCTTACAACTATCTCTTGGCACTGCTGCTGCGGCACTTGTGCTCTTTTCA 307
Qy      241  GTTTTGTGAGCTTCTGTTGGAAGATTTCACTTGAACATGACAGATGCTGCTGCTCC 300
Db      308  GTTTTGTGAGCTTCTGTTGGAAGATTTCACTTGAACATGACAGATGCTGCTGCTCC 367
Qy      301  GACAGATCATAGAAAGTGTGGAATTTCTATCATTCACACCTCTCATATGATTTCTGTA 360
Db      368  GACAGATCATAGAAAGTGTGGAATTTCTATCATTCACACCTCTCATATGATTTCTGTA 427
Qy      361  CGCTTAAACATGACAGATATTCGCTGTGCGACCGGCTCAAGTACACAGGCTCTCA 420
Db      428  CGCTTAAACATGACAGATATTCGCTGTGCGACCGGCTCAAGTACACAGGCTCTCA 487
Qy      421  TACCCAGCCCGACCCGGAAGTCAATGTAAGTTTACATCACTGCTTCTGACACAGC 480
Db      488  TACCCAGCCCGACCCGGAAGTCAATGTAAGTTTACATCACTGCTTCTGACACAGC 547
Qy      481  ATCCCTATTATCTGTTGGCCCAACATCTGAGATGAAAGCTATACAGACCTCTGAGAT 540
Db      548  ATCCCTATTATCTGTTGGCCCAACATCTGAGATGAAAGCTATACAGACCTCTGAGAT 607
Qy      541  CAGTCTCATCTGATGATCACTGCTTCAACCGTCTACCTGATGCTGCTCATCTTCTTC 600

```

Db 608 CAGCTCTCATCTGATGCACTGCTTACCGCTGAGCGCTCATCTTCTTC 667
 Qy 601 ATCTGAACTCATCTGTTGTATACAGTCAAGAGCAATTTTCTCTCGTGC 660
 Db 668 ATCTGAACTCATCTGTTGTATACAGTCAAGAGCAATTTTCTCTCGTGC 727
 Qy 661 TACTCCAGGGGAGAGCAACCGCATTTTGTACCACTTACCTTCATCTTGCACACT 720
 Db 728 TACTCCAGGGGAGAGCAACCGCATTTTGTACCACTTACCTTCATCTTGCACACT 787
 Qy 721 TGGGCCCCCGCATCATCTGATGATTTTACCACTTATGAGGCGCCATCCAGAACGG 780
 Db 788 TGGGCCCCCGCATCATCTGATGATTTTACCACTTATGAGGCGCCATCCAGAACGG 847
 Qy 781 TGGCTGTGACATCATCTGTCGACATTTGCCAATGTTAGCTTCTTGAACACAGCATC 840
 Db 848 TGGCTGTGACATCATCTGTCGACATTTGCCAATGTTAGCTTCTTGAACACAGCATC 907
 Qy 841 AACTCTCTCTACAGCTTCACTGACAGAGGGTTCCGACATGGAGCGGCGACGCTC 900
 Db 908 AACTCTCTCTACAGCTTCACTGACAGAGGGTTCCGACATGGAGCGGCGACGCTC 967
 Qy 901 AAGCTTTCTTCAAGTCCAGAGCAACTGTACAGTTCTACCACTTAACTTTCC 960
 Db 968 AAGCTTTCTTCAAGTCCAGAGCAACTGTACAGTTCTACCACTTAACTTTCC 1027
 Qy 961 ATTAACAAGTACCCCTGATCTGCGCGCAACTCACATGATCAAGATGCTGTGTAC 1020
 Db 1028 ATTAACAAGTACCCCTGATCTGCGCGCAACTCACATGATCAAGATGCTGTGTAC 1087
 Qy 1021 CAGTATGACAAAATGGAATACTTAAAGTATCCCGCTGA 1062
 Db 1088 CAGTATGACAAAATGGAATACTTAAAGTATCCCGCTGA 1129
 RESULT 7
 ID ADF70587 standard; DNA; 1776 BP.
 AC ADF70587;
 DT 12-FEB-2004 (first entry)
 DE Orphan receptor ligand-related human protein gene SeqID210.
 KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KW GFPuv; Enhanced GFP; EGFP; human; gene; ds.
 OS Homo sapiens.
 PN WO2003071272-A1.
 PD 28-AUG-2003.
 PF 21-FEB-2003; 2003WO-JP001901.
 PR 22-FEB-2002; 2002JP-00045728.
 PR 23-JUL-2002; 2002JP-00213949.
 PR 11-OCT-2002; 2002JP-00298237.
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
 DR WPI; 2003-697654/66.
 DR P-PSDB; ADF70485.
 PT Transformation of cells with a fusion protein of an orphan receptor
 PT protein with a fluorescent protein useful for identification of ligands
 PT to the orphan receptor.
 PS Example 4; SEQ ID NO 210; 594bp; Japanese.

XX This invention relates to a novel method of identifying ligands to an
 CC orphan receptor protein which comprises transforming cells with DNA
 CC encoding a fusion protein of the orphan receptor with a fluorescent
 CC protein, so that the fusion protein is expressed in the cells (or cell
 CC membranes isolated from them) and contacting the cells with the potential
 CC ligand to be tested. A suitable fluorescent protein (GFP) for incorporation in
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
 CC identification of ligands binding to an orphan receptor protein.
 SQ Sequence 1776 BP; 471 A; 500 C; 358 G; 447 T; 0 U; 0 Other;
 Query Match 99.8%; Score 1060; DB 10; Length 1776;
 Best Local Similarity 100.0%; Pred. No. 6, 4e-284;
 Matches 1060; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGAGACACACGACCCCACTCGACGCAACAGCTCGTGTGTTGTTCCCGCGC 60
 Db 1 ATGAGACACACGACCCCACTCGACGCAACAGCTCGTGTGTTGTTCCCGCGC 60
 Qy 61 TCGGCTGTGGGCTTGGGTTTGTGCGCGTGTCTACTACAGCTCTTGCTGTGCT 120
 Db 61 TCGGCTGTGGGCTTGGGTTTGTGCGCGTGTCTACTACAGCTCTTGCTGTGCT 120
 Qy 121 TTACACGCAATATCTTGTACAGTATCATCTTCCAGCTGTGTGCAAGAACAGAG 180
 Db 121 TTACACGCAATATCTTGTACAGTATCATCTTCCAGCTGTGTGCAAGAACAGAG 180
 Qy 181 TCTCTACAACTATCTTGTGACATCGGCTGCGGACATCTGTGCTCTTTTTCATA 240
 Db 181 TCTCTACAACTATCTTGTGACATCGGCTGCGGACATCTGTGCTCTTTTTCATA 240
 Qy 241 GTGTGTGTGACCTTCTGTGTGGAAGATTTTATCTTGAACAGATGCTCTAGTCCC 300
 Db 241 GTGTGTGTGACCTTCTGTGTGGAAGATTTTATCTTGAACAGATGCTCTAGTCCC 300
 Qy 301 GACAAATCATAGAAATGCTGTGAAATCTCATCTCAACACCTCCATATGATTAATCTGA 360
 Db 301 GACAAATCATAGAAATGCTGTGAAATCTCATCTCAACACCTCCATATGATTAATCTGA 360
 Qy 361 CGGTAAACATTGACAGATATATCGCTGTGACCCGCTCAAGTACACAGGTCTCA 420
 Db 361 CGGTAAACATTGACAGATATATCGCTGTGACCCGCTCAAGTACACAGGTCTCA 420
 Qy 421 TACCCAGCCCGACCCGGAAGTCAATGTATGTTTACATCACTGCTTCTGACACG 480
 Db 421 TACCCAGCCCGACCCGGAAGTCAATGTATGTTTACATCACTGCTTCTGACACG 480
 Qy 481 ATCCCTATTTATGTTGAGCCCAACATCTGACATGAAGACTATACGACACTTGTGCA 540
 Db 481 ATCCCTATTTATGTTGAGCCCAACATCTGACATGAAGACTATACGACACTTGTGCA 540
 Qy 541 CAGTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Db 541 CAGTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Qy 601 ATCTGAACTCATCTGTTGTATACAGTCAAGAGCAATTTTCTCTCGTGC 660
 Db 601 ATCTGAACTCATCTGTTGTATACAGTCAAGAGCAATTTTCTCTCGTGC 660
 Qy 661 TACTCCAGGGGAGAGCAACCGCATTTTGTACCACTTACCTTCATCTTGCACACT 720
 Db 661 TACTCCAGGGGAGAGCAACCGCATTTTGTACCACTTACCTTCATCTTGCACACT 720
 Qy 721 TGGGCCCCCGCATCATCTGATGATTTTACCACTTATGAGGCGCCATCCAGAACGG 780
 Db 721 TGGGCCCCCGCATCATCTGATGATTTTACCACTTATGAGGCGCCATCCAGAACGG 780
 Qy 781 TGGCTGTGACATCATCTGTCGACATTTGCCAATGTTAGCTTCTTGAACACAGCATC 840
 Db 781 TGGCTGTGACATCATCTGTCGACATTTGCCAATGTTAGCTTCTTGAACACAGCATC 840

QY 841 AACTTCTTCTCTACTGCTTTCATCAGCAGGGGTTCCGACCACTGACGGCCGACGCTC 900
 |||||
 Db 841 AACTTCTTCTCTACTGCTTTCATCAGCAGGGGTTCCGACCACTGACGGCCGACGCTC 900
 |||||
 QY 901 AAGGCTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTTACCAATCAATCACTTTTCC 960
 |||||
 Db 901 AAGGCTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTTACCAATCAATCACTTTTCC 960
 |||||
 QY 961 ATAACAGTAGCCCGGATCTCGCGGCAACTCAGCTGATCAAGATGCTGTGTAC 1020
 |||||
 Db 961 ATAACAGTAGCCCGGATCTCGCGGCAACTCAGCTGATCAAGATGCTGTGTAC 1020
 |||||
 QY 1021 CAGTATGACAAATATGAAATCTATTAAGTATCCCGT 1060
 |||||
 Db 1021 CAGTATGACAAATATGAAATCTATTAAGTATCCCGT 1060
 |||||
 RESULT 8
 ID AAS15731 standard; cDNA, 1343 BP.
 XX AAS15731;
 AC AAS15731;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE DNA encoding chemokine receptor family related protein, NOV11.
 XX
 KM NOV; cytostatic; psoriasis; neutropenic; neuroprotectant;
 KM cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
 KM haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;
 KM differentiation; proliferation; haematopoiesis; wound healing;
 KM angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
 KM haemophilia; allergy; Pendred syndrome; skeletal dysplasia;
 KM ischemic injury; neuroepithelial disorder; hepatitis; heart failure;
 KM chemokine receptor; chromosome 1; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1 /*tag= a
 FT 5'UTR
 FT 2.1063
 FT /*tag= b
 FT CDS /product= "Chemokine receptor related protein, NOV11"
 FT 2.142
 FT sig_peptide /*tag= c
 FT mat_peptide 143..1060
 FT /*tag= d
 FT /note= "Mature chemokine receptor related protein, NOV11"
 FT 1064..1343
 FT /*tag= e
 PN WO200170978-A2.
 PD 27-SEP-2001.
 XX
 PF 20-MAR-2001; 2001WO-US009093.
 XX
 PR 20-MAR-2000; 2000US-0190768P.
 PR 20-MAR-2000; 2000US-0190835P.
 PR 22-MAR-2000; 2000US-0190972P.
 PR 22-MAR-2000; 2000US-0191199P.
 PR 24-MAR-2000; 2000US-0191947P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192664P.
 PR 28-MAR-2000; 2000US-0192665P.
 PR 28-MAR-2000; 2000US-0192984P.
 PR 29-MAR-2000; 2000US-0192835P.
 PR 31-MAR-2000; 2000US-0193843P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Taupier RJ, Majumder K, Spaderma SK, Smithson G, Mezes PS;

PI Vernet CAM;
 XX
 DR MPI; 2001-639127/73.
 DR P-PSDB; AAU10068.
 XX
 PT Polypeptides and nucleic acids related to chloride channel, insulin-like
 PT growth factor family of proteins, useful for diagnosing and treating
 PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
 XX
 PS Claim 9; Page 44; 151pp; English.
 XX
 CC The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
 CC polypeptides are useful for treating pathology associated with NOVX
 CC polypeptide, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX, identifying agents binding to
 CC NOVX and treatment of disorders associated with altered expression of
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and
 CC congenital myotonia. NOVX proteins are useful in treatment of disorders
 CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
 CC in one or more organs (e.g. haemophilia, anaemia), Pendred syndrome,
 CC skeletal dysplasias, disorders characterised by altered cell shape,
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
 CC treatment of disorders of vascular smooth muscle cell differentiation,
 CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
 CC useful to screen for molecules which inhibit or enhance NOVX activity or
 CC function and are useful as targets for the identifying small molecules,
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
 CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV
 CC sequences are also useful for: identifying a cell or tissue type in a
 CC biological sample; amplifying DNA sequences from very small biological
 CC samples e.g. hair or skin or body fluids and as primers and probes to
 CC identify and/or clone NOVX homologues. NOVX proteins are useful
 CC immunogens to generate antibodies to monitor protein levels and modulate
 CC NOVX activity. Cells comprising the nucleic acids are useful for
 CC producing transgenic animals, for studying the function and/or activity
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX
 CC protein activity. This sequence encodes NOV11 (located on chromosome 1)
 CC related to the chemokine receptor family of proteins, one of 12 NOV
 CC polypeptides described in the method of the invention
 XX
 SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
 Query Match 99.7%; Score 1058.8; DB 4; Length 1343;
 Best Local Similarity 99.8%; Pred. No. 1.2e-283;
 Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGAGCAGCAGCAGCGCCACTCTGCAAGCCAAAGCTGCTGTGTGGTGTCGCCGCG 60
 Db 2 ATGAGCAGCAGCAGCGCCACTCTGCAAGCTGCTGTGTGGTGTCGCCGCG 61
 QY 61 TCGGCTGCGGCTTGGGTTTCTGTCGCCGTGCTTACTACAGCTCTTGTGCTGCTGGT 120
 Db 62 TCGGCTGCGGCTTGGGTTTCTGTCGCCGTGCTTACTACAGCTCTTGTGCTGCTGGT 121
 QY 121 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGGAAGAAGCAGAG 180
 Db 122 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGGAAGAAGCAGAG 181
 QY 181 TCTCTTACAACTATCTCTTGGCACTGCTGCTGCCAGATCTTGTCTCTTTTTCATA 240
 Db 182 TCTCTTACAACTATCTCTTGGCACTGCTGCTGCCAGATCTTGTCTCTTTTTCATA 241
 QY 241 GTGTTTGTGACTTCTTCTTGGGAAGATTTTCACTTGAACATGCAAGATGCTCAGGTCCC 300
 Db 242 GTGTTTGTGACTTCTTCTTGGGAAGATTTTCACTTGAACATGCAAGATGCTCAGGTCCC 301
 QY 301 GACAAGATCATAGAAGTGTGAATTTCTCATTCATCCAGACCTTCATATGATTAATCTGTA 360
 Db 302 GACAAGATCATAGAAGTGTGAATTTCTCATTCATCCAGACCTTCATATGATTAATCTGTA 361
 QY 361 CCGTTAACATGACAGTATATGCTGTCTGCGACCCGCTCAAGTACACACGGTCTCA 420

[illegible][illegible]

Db 422 TACCCAGCCCGACCGGAAAGTCATGTAAGTGTTCATCATCTGCTTCCAGCAGC 481
Qy 481 ATCCCTATTAATCTGGTGGCCCAACATCTGAGCTGAAGATCATAGCAGCCTGTGAT 540
Db 482 ATCCCTATTAATCTGGTGGCCCAACATCTGAGCTGAAGATCATAGCAGCCTGTGAT 541
Qy 541 CAGCTCTCATCTGATCATCTGCTTCAACCGCTTACCTGGTGGCCCTGTGATCTTCTTC 600
Db 542 CAGCTCTCATCTGATCATCTGCTTCAACCGCTTACCTGGTGGCCCTGTGATCTTCTTC 601
Qy 601 ATCTTGAATCATATATTTGTGTACAGCTCAGAGAGAGACAAATTTTCTGCTCGAGGC 660
Db 602 ATCTTGAATCATATATTTGTGTACAGCTCAGAGAGAGACAAATTTTCTGCTCGAGGC 661
Qy 661 TACTCCAGGGGAGACACCGGCACTTGTTCACATTAATCTCATCTTTGCAACCTT 720
Db 662 TACTCCAGGGGAGACACCGGCACTTGTTCACATTAATCTCATCTTTGCAACCTT 721
Qy 721 TGGGCCCCCGCATCATCATATTTTACCACTTATGGGGGCGCCATTCAGAACCGC 780
Db 722 TGGGCCCCCGCATCATCATATTTTACCACTTATGGGGGCGCCATTCAGAACCGC 781
Qy 781 TGGCTGGTGCATCATCATCTGCTTCAACAGCTGCTAGCCCTTCTGAACAGAGCATC 840
Db 782 TGGCTGGTGCATCATCATCTGCTTCAACAGCTGCTAGCCCTTCTGAACAGAGCATC 841
Qy 841 AACTCTTCTCTCATCTGCTTCAATCAGCAGAGCGGTTCCGACATTCAGAGCGGCGCTC 900
Db 842 AACTCTTCTCTCATCTGCTTCAATCAGCAGAGCGGTTCCGACATTCAGAGCGGCGCTC 901
Qy 901 AAGGCTTTCTTCAAGTGCAGAGACAACTGTACAGTTCTTACACATCATTAATTTTCC 960
Db 902 AAGGCTTTCTTCAAGTGCAGAGACAACTGTACAGTTCTTACACATCATTAATTTTCC 961
Qy 961 ATACAGAGTACCCCTGATCTGCGCGGCAACTCATCTGATCATGAAGTCTGTGTAC 1020
Db 962 ATACAGAGTACCCCTGATCTGCGCGGCAACTCATCTGATCATGAAGTCTGTGTAC 1021
Qy 1021 CAGTATGACAAATATGAAAACTTATTAATATCCCGTGA 1062
Db 1022 CAGTATGACAAATATGAAAACTTATTAATATCCCGTGA 1063

RESULT 10
AD179324
ID AD179324 standard; DNA; 1343 BP.
XX
AC AD179324;
XX

22-APR-2004 (first entry)
XX
DE NOVI1 coding sequence, SEQ ID 21.
XX

XX Cytostatic; Immunosuppressive; Antiallergic; Antimicrobial; Vasoactive;
XX Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
XX Ophthalmological; Antiparasitic; Neuroprotective; Nocitropic;
XX Antileukosclerotic; Hypotensive; Cardiac; Cerebroprotective;
XX Gene Therapy; NOVI1; human; cancer; myelogenous leukemia;
XX allergic; infectious; asthma; lung disease; reproductive disorder;
XX haemangioma; death; liver cirrhosis; hepatitis C; gastric disorder;
XX diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;
XX hypertension; stroke; heart failure; chromosome 1; NOVI1;
XX chemokine receptor; gene; ds.

OS Homo sapiens.
XX

XX Key Location/Qualifiers
FH 2..1063
FT /tag= a
FT /product= "NOVI1"
XX

PN WO2004009635-A2.
XX
PD 29-JAN-2004.
XX
PF 04-OCT-2001; 2001WO-US031292.
XX
PR 20-MAR-2001; 2001US-00813432.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Traupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;
PI Verne C;
XX
DR MPI; 2004-123380/12.
XX
DR P-PSDB; AD179325.
XX
PS Claim 8; Page 43-44; 158pp; English.

CC The present invention relates to novel NOVI proteins and their coding
CC sequences (AD179304-AD179327). The sequences are useful for the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease associated with the protein, or for diagnosing and treating
CC disorders associated with the NOVI protein, such as cancer, myelogenous
CC leukaemia, congenital neonatal autoimmune thrombocytopenia,
CC immunological disorders, allergy and infection, asthma, lung diseases,
CC reproductive disorders, male and female reproductive diseases,
CC haemangioma, death, liver cirrhosis, hepatitis C, gastric disorders,
CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
CC hypertension, stroke and heart failure. NOVI1 represents a new member of
CC the chemokine receptor family and the gene is located on chromosome 1.
XX
SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Query Match 99.7%; Score 1058.8; DB 12; Length 1343;
Best Local Similarity 99.8%; Pred. No. 1.2e-283;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGAGACACGACGCGCCACTCTGACGACCAAGCTGCTGTGATGATCCCGGCG 60
Db 2 ATGAGAGACACGACGCGCCACTCTGACGACCAAGCTGCTGTGATGATCCCGGCG 61
Qy 61 TCGGCTGCGGCTTGGGTTTCGCGCGGTGCTACTACAGCTCTTGTGCTGCTCGGT 120
Db 62 TCGGCTGCGGCTTGGGTTTCGCGCGGTGCTACTACAGCTCTTGTGCTGCTCGGT 121
Qy 121 TTACACAGCAATATCTTGAACAGTATCATCTTCCAGCTGTGACAGAGAGAG 180
Db 122 TTACACAGCAATATCTTGAACAGTATCATCTTCCAGCTGTGACAGAGAGAG 181
Qy 181 TCCCTCCACAACTATCTTTGGACATCGGCTGCTGCGACATTTGGTCTCTTTTCATA 240
Db 182 TCCCTCCACAACTATCTTTGGACATCGGCTGCTGCGACATTTGGTCTCTTTTCATA 241
Qy 241 GTGTTTGTGACCTCTCTGTGAGATTTCACTTGAACATGACAGTCTCAGGTCCG 300
Db 242 GTGTTTGTGACCTCTCTGTGAGATTTCACTTGAACATGACAGTCTCAGGTCCG 301
Qy 301 GACAGATATATAGTGTGAGATTTCACTTCAATCCACACCTTCATATGATTA 360
Db 302 GACAGATATATAGTGTGAGATTTCACTTCAATCCACACCTTCATATGATTA 361
Qy 361 CCGTTAACCATTTGACAGTATATGCTGTGACACCGGCTCAGTACCAAGGTCTCA 420
Db 362 CCGTTAACCATTTGACAGTATATCTGTGACACCGGCTCAGTACCAAGGTCTCA 421
Qy 421 TACCCAGCCCGACCGGAAAGTCAATGTAAGTGTATACATCACTGCTTCTGACGAG 480
Db 422 TACCCAGCCCGACCGGAAAGTCAATGTAAGTGTATACATCACTGCTTCTGACGAG 481
Qy 481 ATCCCTATTAATCTGGTGGCCCAACATCTGAGCTGAAGATCATAGCAGCCTGTGAT 540
Db 482 ATCCCTATTAATCTGGTGGCCCAACATCTGAGCTGAAGATCATAGCAGCCTGTGAT 541

QY 661 TACTCCAGGGGGAAGACACCGCCATCTTGTTCACCTTACCTTCATCTTGTCCACACTT 720
 DB 1113 TACTCCAGGGGGAAGACACCGCCATCTTGTTCACCTTACCTTGTCCACACTT 1172
 QY 721 TGGGCCCCCGGCATCATCATGATTTCTTTACCACTCTATGAGGGGCGCCATCCAGAACCG 780
 DB 1173 TGGGCCCCCGGCATCATCATGATTTCTTTACCACTCTATGAGGGGCGCCATCCAGAACCG 1232
 QY 781 TGGCTGTGACATCATGTCGCGACATTGCCAACATGCTAGCCCTTGTGAACACAGCCATC 840
 DB 1233 TGGCTGTGACATCATGTCGCGACATTGCCAACATGCTAGCCCTTGTGAACACAGCCATC 1292
 QY 841 AACTTCTTCTCTACCTGCTTATCATCAGAACCGGTTCCGCAACATGAGAGCCGCGACGCTC 900
 DB 1293 AACTTCTTCTCTACCTGCTTATCATCAGAACCGGTTCCGCAACATGAGAGCCGCGACGCTC 1352
 QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACACCAATCATTAATTCTTTC 960
 DB 1353 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACACCAATCATTAATTCTTTC 1412
 QY 961 ATACAGATGAGCCCTGATCTTCCGCGCAACTCAGCTGATCAAGATGCTGTGTAC 1020
 DB 1413 ATACAGATGAGCCCTGATCTTCCGCGCAACTCAGCTGATCAAGATGCTGTGTAC 1472
 QY 1021 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGCTGA 1062
 DB 1473 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGCTGA 1514

RESULT 13

ABN84269
 ID ABN84269 standard; cDNA; 1059 BP.

AC ABN84269;

XX 23-SEP-2002 (first entry)

DT Human chemokine-like receptor cDNA.

XX Chemokine-like receptor; G-protein coupled receptor; receptor; human;

KW HIV infection; cardiovascular disease; asthma;

KW chronic obstructive pulmonary disease; cardiant; antiaesthetic;

KW vasotropic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;

KW antiinflammatory; antiallergic; immunomodulator; gene therapy; gene; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1059

FT /tag= a

FT /product= "Chemokine-like receptor"

FT /note= "the CDS does not include a stop codon"

XX MO200248358-A2.

XX 20-JUN-2002.

XX 12-DEC-2001; 2001WO-EP014571.

XX 14-DEC-2000; 2000US-0255150P.

XX 02-APR-2001; 2001US-0280110P.

XX 21-JUN-2001; 2001US-0299474P.

XX (FARB) BAYER AG.

XX Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;

XX WPI; 2002-547858/58.

XX P-PSDB; ABB79517.

XX New isolated polynucleotide encoding a chemokine-like receptor

PT polypeptide for treating e.g. asthma, myocardial infarction, human
 PT immunodeficiency virus infection, or chronic obstructive pulmonary
 PT disease.

PS Claim 1; Fig 1; 114pp; English.

CC The present sequence is that of cDNA encoding a novel human chemokine-
 CC like receptor (see ANN79517) of 353 amino acids and having 7 putative
 CC transmembrane domains, consistent with the structure of a G-protein
 CC coupled receptor. Its closest human homologue is C-C chemokine receptor
 CC 3. The novel receptor is expressed at low levels in most tissues. It is
 CC expressed at a high level in phytohemagglutinin-stimulated CD8+ cells,
 CC but in none of the other immune cells tested. It may act as a receptor of
 CC chemottractant molecules on activated lymphocytes and be involved in
 CC cell trafficking and homing to sites of infection, inflammation or tissue
 CC injury. Regulation of activity of the novel receptor can therefore be
 CC used to treat cardiovascular, immunological and inflammatory diseases,
 CC including asthma and chronic obstructive pulmonary disease (COPD). The
 CC receptor may also be a target for viruses that reside in the nervous
 CC system. Regulating the binding of ligands, e.g. chemottractant molecules
 CC or virus particles, to the receptor can therefore be used to modulate the
 CC immune response to inhibit viral infections, including HIV infection. A
 CC claimed method of reducing activity of the receptor involves contacting a
 CC cell with a reagent (preferably an antibody, antisense oligonucleotide or
 CC ribozyme) to a product (preferably RNA or a polypeptide) encoded by a
 CC polynucleotide encoding the human chemokine-like receptor in vivo or in
 CC vitro. A claimed method of treating a chemokine-like receptor dysfunction
 CC related disease selected from HIV infection, a cardiovascular disorder,
 CC asthma or COPD uses a reagent that modulates a function of the receptor
 CC XX

SO Sequence 1059 BP; 240 A; 342 C; 206 G; 271 T; 0 U; 0 Other;

Query Match 99.6%; Score 1057.4; DB 6; Length 1059;

Best Local Similarity 99.9%; Pred. No. 2,6e-283;

Matches 1058; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGACCCACCTCGACGCAACGCTGCTGTGAGTCCCGGCG 60
 DB 1 ATGAGACACACGACGACCCACCTCGACGCAACGCTGCTGTGAGTCCCGGCG 60
 QY 61 TCGGCTGCGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGTGCTCGGT 120
 DB 61 TCGGCTGCGGCTTGGGTTTGGCCGCTGCTACTACAGCTCTTGTGCTCGGT 120
 QY 121 TTACACGCAAAATATCTTGAACATGATCTCTCCAGCTGTGCAAGACGAA 180
 DB 121 TTACACGCAAAATATCTTGAACATGATCTCTCCAGCTGTGCAAGACGAA 180
 QY 181 TCCTCTAACATATCTTGGACATCGCTGCGCGGACATGTGGTCTCTTTTCAT 240
 DB 181 TCCTCTAACATATCTTGGACATCGCTGCGCGGACATGTGGTCTCTTTTCAT 240
 QY 241 GTGTTTGTGACCTTCTGTGGAAGATTTCATCTTGAACATGCTGAGTCCCG 300
 DB 241 GTGTTTGTGACCTTCTGTGGAAGATTTCATCTTGAACATGCTGAGTCCCG 300
 QY 301 GACAGATCATAGAGTGTGGAATTTCTATCCATCCACACCTCCATATGATTTCT 360
 DB 301 GACAGATCATAGAGTGTGGAATTTCTATCCATCCACACCTCCATATGATTTCT 360
 QY 361 CGGTTAACATGACAGATATCGCTGTCGACCGGCTCAAGTACACACGCTCTCA 420
 DB 361 CGGTTAACATGACAGATATCGCTGTCGACCGGCTCAAGTACACACGCTCTCA 420
 QY 421 TACCCAGCCGACCCGGAAGTCAATTTAGTGTTTTACATCACTGCTTCTGACAC 480
 DB 421 TACCCAGCCGACCCGGAAGTCAATTTAGTGTTTTACATCACTGCTTCTGACAC 480
 QY 481 ATCCCTTATTAATGTTGGTCCCAACATCTGATGGAAGATTAATACAGCACTTGT 540
 DB 481 ATCCCTTATTAATGTTGGTCCCAACATCTGATGGAAGATTAATACAGCACTTGT 540
 QY 541 CAGCTCTATCTGATCCATGCTTACCGCTTACCGTCTAGTCCGCTGCATCTTCT 600

Db	541	CACGCTCCATCTGTGGATCCACTGGTTCACCGCTTAACTGGTGGCCGTGCTCAATCTTCTTC	600
Qy	601	ATCTTGAACATCAATCATTTGTGTACAGCTCAGAGAGAAAGCAATTTTGTCTTCGGTGGC	660
Db	601	ATCTTGAACATCAATCATTTGTGTACAGCTCAGAGAGAAAGCAATTTTGTCTTCGGTGGC	660
Qy	661	TACTCCACGGGGAAAGACCAACCGGCATCTTGTTCACATTAACCTTCATCTTTGGCAGACTT	720
Db	661	TACTCCACGGGGAAAGACCAACCGGCATCTTGTTCACATTAACCTTCATCTTTGGCAGACTT	720
Qy	721	TGGGCCCCCGGCATCATCATGATATTTCTTTACACCTCTATAGGGGGCCGCATCCAGAACGGC	780
Db	721	TGGGCCCCCGGCATCATCATGATATTTCTTTACACCTCTATAGGGGGCCGCATCCAGAACGGC	780
Qy	781	TGGCTGGTGCACTCATATGTCGCAATATGCCAATCATCTGACCTTCTGAAACAAGCATTC	840
Db	781	TGGCTGGTGACATCATATGTCGCAATATGCCAATCATCTGACCTTCTGAAACAAGCATTC	840
Qy	841	AACCTTCCTCCCTACAGCTTTCATCAGCAAGCGGTTCCGACCATATGGCAGCGCCACGCTC	900
Db	841	AACCTTCCTCCCTACAGCTTTCATCAGCAAGCGGTTCCGACCATATGGCAGCGCCACGCTC	900
Qy	901	AAGGCTTTTCTTCAAGTGCAGAAAGCAACCTGTACAGTTTCAACCAATCATTAATCTTTTCC	960
Db	901	AAGGCTTTTCTTCAAGTGCAGAAAGCAACCTGTACAGTTTCAACCAATCATTAATCTTTTCC	960
Qy	961	ATPACAAAGAGCGCCCGATCTCGCGGGAAATCTCACTGATCAAGATGTGGTGTAC	1020
Db	961	ATPACAAAGAGCGCCCGATCTCGCGGGAAATCTCACTGATCAAGATGTGGTGTAC	1020
Qy	1021	CAGTATGACAAAATGGAAGAAACCTTATTAAGATATCCCGC	1059
Db	1021	CAGTATGACAAAATGGAAGAAACCTTATTAAGATATCCCGC	1059
RESULT 14			
AAD46858			
ID	AAD46858 standard; cDNA; 1526 BP.		
XX	AAD46858;		
AC	27-JAN-2003 (first entry)		
DT			
XX			
DE	Human 7TM domain receptor 65507 cDNA.		
XX			
KW	Human; adenosine deaminase; seven transmembrane domain receptor; cancer;		
KW	7TM; glycoprotease; immune disorder; IGA deficiency; allergy; arrhythmia;		
KW	rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;		
KW	hypertension; ischaemic heart disease; obesity; myocardial infarction;		
KW	endothelial cell disorder; Graves' disease; psoriasis; brain disorder;		
KW	Parkinson's disease; Alzheimer's disease; haematopoietic disorder;		
KW	cerebral oedema; metabolic disorder; liver disorder; platelet disorder;		
KW	chromosome mapping; tissue typing; gene therapy; neuroprotective;		
KW	cytostatic; anorectic; cardiant; haemostatic; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	139..1200		
FT	/*tag= a		
FT	/product= "Human 65507 protein"		
FT	/note= "This region is specifically claimed as SEQ ID NO:		
FT	12 in claim 1 of the specification"		
FT	139..306		
FT	/*tag= b		
FT	mat_peptide		
FT	307..1197		
FT	/*tag= c		
FT	/product= "Human mature 65507 protein"		
XX			
PN	WO200274960-A2.		
XX			
DD	26-SEP-2002.		

XX	PF	08-NOV-2001; 2001WO-US051427.
XX	PR	08-NOV-2000; 2000US-0246768P.
XX	PR	08-NOV-2000; 2000US-0246772P.
XX	PR	15-NOV-2000; 2000US-0249185P.
XX	PA	(MILL-) MILLENNIUM PHARM INC.
XX	PI	Leidy KR, Kapeller-Libermann R, Glucksmann M;
XX	DR	WPI, 2002-759898/82.
XX	PS	P-PSDB; AAE29236.
PT	PT	New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,
PT	PT	useful for diagnosing and treating cancer, immune, cardiovascular,
PT	PT	hematopoietic, brain, pain, metabolic, liver or platelet disorders, and
XX	XX	in pharmacogenomics.
XX	PS	Claim 1; Fig 17; 178pp; English.
CC	CC	The present invention relates to novel 38650, 28472, 5495, 65507, 81588
CC	CC	or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-
CC	CC	protease or seven transmembrane domain (7TM) receptor family members.
CC	CC	Sequences of the invention are useful in diagnosing and treating cancer
CC	CC	or aberrant cellular proliferation and/or differentiation (e.g. colon or
CC	CC	lung cancer), immune disorders (e.g. selective IgA deficiency, rheumatoid
CC	CC	arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g.,
CC	CC	hypertension, atherosclerosis, arrhythmias, ischemic heart disease, e.g.,
CC	CC	myocardial infarction, thrombus) including endothelial cell disorders
CC	CC	(e.g. psoriasis, Grave's disease), hematopoietic disorders, brain
CC	CC	disorders (e.g. cerebral oedema, Parkinson's or Alzheimer's disease),
CC	CC	pain and metabolic disorders (e.g. obesity), liver disorders or platelet
CC	CC	disorders. They are also useful in screening assays, predictive medicine
CC	CC	(e.g. diagnostic assays, prognostic assays), monitoring clinical trials
CC	CC	and pharmacogenetics) and prophylactic and therapeutic methods. The
CC	CC	nucleic acids may also be used in chromosome mapping, tissue typing and
CC	CC	forensic biology and as surrogate markers. Sequences of the invention are
CC	CC	also used in gene therapy. The present sequence is human 7TM domain
XX	XX	receptor 65507 cDNA
XX	XX	Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
XX	XX	Query Match 98.9%; Score 1050.8; DB 6; Length 1526;
XX	XX	Best Local Similarity 99.3%; Pred. No. 2.1e-281;
XX	XX	Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	QY	1 ATGGAGCACAGCAGCCCACTTGAGCCCAAGCTCGCTGTCCTTGTCCTCCCGGAC 60
DB	DB	139 ATGGAGCACAGCAGCCCACTTGAGCCCAAGCTCGCTGTCCTTGTCCTCCCGGAC 198
QY	QY	61 TCGGCGCTGGGCTGGGTTTCGGCCCGTGTCTATCAAGCTCTTGCTGTGCTCGGT 120
DB	DB	199 TCGGCGCTGGGCTGGGTTTCGGCCCGTGTCTATCAAGCTCTTGGAAGGCTCGGT 258
QY	QY	121 TTACCAAGAAATATCTTGACAGTATCATCTCTCCAGCTGTGGCAAGACAGAAAG 180
DB	DB	259 TTACCAAGAAATATCTTGACAGTATCATCTCTCCAGCTGTGGCAAGACAGAAAG 318
QY	QY	181 TCCCTCTACATCTCTCTTGGAAGCTCGCTGCTGCCGAATCTTGCTCTCTTTTCATA 240
DB	DB	319 TCCCTCTACATCTCTCTTGGAAGCTCGCTGCTGCCGAATCTTGCTCTCTTTTCATA 378
QY	QY	241 GTGTTTGGACCTTCTGTGGAAGATTCATCTTGAAATGACATGATGCTCAGGTCCC 300
DB	DB	379 GTGTTTGGACCTTCTGTGGAAGATTCATCTTGAAATGACATGATGCTCAGGTCCC 438
QY	QY	301 GACAAGATCATAGAAAGCTGGAATTCATCATATCCACACTCCATATGATTAATCTGA 360
DB	DB	439 GACAAGATCATAGAAAGCTGGAATTCATCATATCCACACTCCATATGATTAATCTGA 498
QY	QY	351 CCGTTAACCATGACAGTATATCGCTGTCTCCACCGCTCAAGTACCAAGGATCTCA 420

Db 499 CCCTTAACATGTAGACAGTATATCGCTGTGCGACCCCGCTCAAGTACACAGGCTCTCA 558
 QY 421 TACCCAGCCCGCACCCGGAAAGTCAATTGTTAATCACTGCTTCTGACAGC 480
 Db 559 TACCCAGCCCGCACCCGGAAAGTCAATTGTTAATCACTGCTTCTGACAGC 618
 QY 481 ATCCCTATTTACGTGAGGCGCCCAATCTGACGTGAAGACATCACTGACCTCTGAGCAT 540
 Db 619 ATCCCTATTTACGTGAGGCGCCCAATCTGACGTGAAGACATCACTGACCTCTGAGCAT 678
 QY 541 CACGCTCTCATCTGATATCCACTGCTTCAACGCTTACCTGAGCTGCTCATCTTCTTC 600
 Db 679 CAGCTCTCATCTGAGTCCACTGCTTCAACGCTTACAGGCTGCTGCTCATCTTCTTC 738
 QY 601 ATCTTGAATCAATCAATTGTGTAACAAGCTCAGAGAGAGAAATTTTCTGCTCGTGGC 660
 Db 739 ATCTTGAATCAATCAATTGTGTAACAAGCTCAGAGAGAGAAATTTTCTGCTCGTGGC 798
 QY 661 TACTCAGCGGGAGAGACACCGGCACTTGTGACCAATTCCTCCATCTTTGCGACACTT 720
 Db 799 TACTCAGCGGGAGAGACACCGGCACTTGTGACCAATTCCTCCATCTTTGCGACACTT 858
 QY 721 TGGGCCCCCGCATCATCATGATTTCTTTACACCTATGAGGCGCCATCCAGAACCGC 780
 Db 859 TGGGCCCCCGCATCATCATGATTTCTTTACACCTATGAGGCGCCATCCAGAACCGC 918
 QY 781 TGGCTGTGACATCATGATGTCGACATTTGCCAATGCTAGCCCTTTGAAACAGACCATC 840
 Db 919 TGGCTGTGACATCATGATGTCGACATTTGCCAATGCTAGCCCTTTGAAACAGACCATC 978
 QY 841 AACTTTCTCTCTCAATGCTTCAATCAGACAGCGGTTCCGACCAATGACAGCGGCGACAGC 900
 Db 979 AACTTTCTCTCTCAATGCTTCAATCAGACAGCGGTTCCGACCAATGACAGCGGCGACAGC 1038
 QY 901 AAGGCTTTCTTCAAGTGCAGAGAACCACTGTACAGTTCTTACACCAATCAATCTTTTCC 960
 Db 1039 AAGGCTTTCTTCAAGTGCAGAGAACCACTGTACAGTTCTTACACCAATCAATCTTTTCC 1098
 QY 961 ATTAACAGTAGACCCCTGAGATCTTCCGCGGCAACTCACTGATCAAGATGCTGTGTAC 1020
 Db 1099 ATTAACAGTAGACCCCTGAGATCTTCCGCGGCAACTCACTGATCAAGATGCTGTGTAC 1158
 QY 1021 CAGTAGACAAAATGGAATACTATTAAGTATCCCGCGTA 1062
 Db 1159 CAGTAGACAAAATGGAATACTATTAAGTATCCCGCGTA 1200

RESULT 15
 ACA60889
 ID ACA60889 standard, cDNA; 1526 BP.
 XX
 AC ACA60889;
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DE Human cDNA encoding 65507 protein.
 XX
 XX Human; sex; gene; cancer; aberrant cellular proliferation;
 KM differentiation; immune disorders; heart disorder; brain disorder;
 KM cardiovascular disorder; endothelial cell disorder; pain disorder;
 KM haemopoietic disorder; blood vessel disorder; metabolic disorder;
 KM liver disorder; platelet disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 139..1200
 FT /tag= a
 FT /product= "Protein 65507"
 FT /note= "This CDS is specifically claimed in claim 1"
 XX
 US2003009017-A1.

PD 09-JUN-2003.
 XX
 PF 08-NOV-2001; 2001US-00012140.
 XX
 PR 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX
 PA (LEIB/) LEIBY K R.
 PA (KAP/) KAPPELLER-LIBERMANN R.
 PA (GLUC/) GLUCKSMANN M A.
 PI
 XX
 XX Leiby KR, Kappeller-Libermann R, Glucksmann MA;
 DR WPI; 2003-428888/40.
 DR P-PSDB; AB09571.
 XX
 PT New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
 PT molecule, useful for diagnosing, treating cancer, pain, or immune,
 PT heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
 PT and liver disorders.
 PS
 XX
 PS Claim 2; Fig 17; 90pp; English.
 CC The invention relates to an isolated 38650 (encoding adenosine
 CC deaminase), 28472 (encoding a glycoprotease), 5495 (encoding a 7
 CC transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
 CC a sequence which is at least 60% identical to the six nucleic acids or
 CC their open reading frames, fragments of at least 15 nucleotides,
 CC naturally occurring variants, or a DNA insert of the plasmid deposited
 CC with the American Type Culture Collection as Accession No. not defined in
 CC the specification, which encodes the nucleic acids (used to produce the
 CC are a host cell containing the nucleic acids (used to produce the
 CC proteins), the encoded proteins, an antibody that selectively binds to
 CC the polypeptide, and identifying a compound that binds to/modulates the
 CC activity of the polypeptide. The nucleic acid molecules, polypeptides and
 CC methods are useful for diagnosing, treating cancer, aberrant cellular
 CC proliferation and/or differentiation, immune disorders, heart disorders,
 CC cardiovascular disorders including endothelial cell disorders,
 CC haemopoietic disorders, blood vessel disorder, brain disorders, pain
 CC and metabolic disorders, liver disorders and platelet disorders (many
 CC examples of these disorders are given in the specification). The present
 CC sequence is the Human cDNA 65507
 CC
 XX
 SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 Query Match 98.9%; Score 1050.8; DB 8; Length 1526;
 Best Local Similarity 99.3%; Pred. No. 2,1e-281;
 Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATGAGACACACGACGCCCACTTCGACGCAACAGCTGCTGTGTTGTTCCCGCCG 60
 Db 139 ATGAGACACACGACGCCCACTTCGACGCAACAGCTGCTGTGTTGTTCCCGCCG 198
 QY 61 TCGGCTGCGGCTTGGGTTTGGTCCCGGTGCTCTAAGGCTCTTGGTGGCTCGGT 120
 Db 199 TCGGCTGCGGCTTGGGTTTGGTCCCGGTGCTCTAAGGCTCTTGGTGGCTCGGT 258
 QY 121 TTACACGAAATATCTTGAACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAAG 180
 Db 259 TTACACGAAATATCTTGAACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAAG 318
 QY 181 TCTCTCAACATATCTTGGCACTGCTGCTGCCGACATCTTGGTCTCTTTTCA 240
 Db 319 TCTCTCAACATATCTTGGCACTGCTGCTGCCGACATCTTGGTCTCTTTTCA 378
 QY 241 GTTTTGGAGCTTCTGTTGGAAGATTCACTTGAACAGCAATGCTCAGGCTCGCC 300
 Db 379 GTTTTGGAGCTTCTGTTGGAAGATTCACTTGAACAGCAATGCTCAGGCTCGCC 438
 QY 301 GACAAGATATGAAGTGTGGAATTTCTCATCATCCACACCTCATATGATTA 360
 Db 439 GACAAGATATGAAGTGTGGAATTTCTCATCATCCACACCTCATATGATTA 498

```
QY 361 CCGTTACCATGACAGGTATATGCTGTGTCACCCGCTCAAGTACACACGGGTCTCA 420
Db |||||
QY 499 CCGTTAACCATGACAGGTATATCGCTGTGTCACCCGCTCAAGTACACACGGGTCTCA 558
Db |||||
QY 421 TACCAGCCCGCAGCCCGGAAAGTCATTTAGTGTTTTATCACCCTGTTCTTGACCAAGC 480
Db |||||
QY 559 TACCAGCCCGCAGCCCGGAAAGTCATTTAGTGTTTTATCACCCTGTTCTTGACCAAGC 618
QY 481 ATCCCTTATTAAGTGTGAGCCCAACATCTGACCTGAGAGACTATCAGACACCTGTGAT 540
Db |||||
QY 619 ATCCCTTATTAAGTGTGAGCCCAACATCTGACCTGAGAGACTATCAGACACCTGTGAT 678
QY 541 CACGTCCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db |||||
QY 679 CACGTCCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 738
QY 601 ATCTTGAACTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db |||||
QY 739 ATCTTGAACTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 798
QY 661 TACTCCACGGGGAAGACACCGGCATCTTGTTCACATTAACCTTCATCTTGTGACACACTT 720
Db |||||
QY 799 TACTCCACGGGGAAGACACCGGCATCTTGTTCACATTAACCTTCATCTTGTGACACACTT 858
QY 721 TGGGCCCCCGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db |||||
QY 859 TGGGCCCCCGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 918
QY 781 TGGGTGTGACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db |||||
QY 919 TGGGTGTGACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978
QY 841 AACTTCTTCTTACTGCTTCATCAGCAAGCGGTTCGACCAATGGCAGCCGCAAGCTTC 900
Db |||||
QY 979 AACTTCTTCTTACTGCTTCATCAGCAAGCGGTTCGACCAATGGCAGCCGCAAGCTTC 1038
QY 901 AAGGCTTCTTCAAGTGTGAGCAACCTGTACAGTTCTACACCAATCATTAACCTTTTCC 960
Db |||||
QY 1039 AAGGCTTCTTCAAGTGTGAGCAACCTGTACAGTTCTACACCAATCATTAACCTTTTCC 1098
QY 961 ATAAACAAGTAGCCCTGGAATCTGCGGGGAAACTGACACTGCATCAAGATGCTGTGTAC 1020
Db |||||
QY 1099 ATAAACAAGTAGCCCTGGAATCTGCGGGGAAACTGACACTGCATCAAGATGCTGTGTAC 1158
QY 1021 CAGTATGACAAAATGAAAAAAGCTATATAAGTATCCCGGTGA 1062
Db |||||
QY 1159 CAGTATGACAAAATGAAAAAAGCTATATAAGTATCCCGGTGA 1200
```

Search completed: December 14, 2004, 22:43:54
Job time : 500.373 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:27:19 ; Search time 89.3365 Seconds
(without alignments)
8449.603 Million cell updates/sec

Title: US-09-995-225B-15

Perfect score: 1062

Sequence: 1 atggagcagacagcagccca.....ctataaagatccccgrrga 1062

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*\n2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*\n3: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*\n4: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*\n5: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*\n6: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	63.2	6.0	10304	4 US-09-627-465B-1	Sequence 1, App11
C 2	59.6	5.6	7218	1 US-08-232-463-14	Sequence 14, App1
C 3	46.8	4.4	2518	3 US-09-433-699-3	Sequence 3, App11
4	46.2	4.4	951	1 US-08-671-525B-1	Sequence 1, App11
5	46.2	4.4	951	1 US-08-672-109B-1	Sequence 1, App11
6	46.2	4.4	951	1 US-08-842-045-1	Sequence 1, App11
7	46.2	4.4	951	2 US-08-842-238-1	Sequence 1, App11
8	46.2	4.4	951	3 US-08-629-335B-1	Sequence 1, App11
9	46.2	4.4	1149	4 US-09-868-552-1	Sequence 1, App11
10	46.2	4.4	1149	4 US-09-868-552-3	Sequence 3, App11
11	46.2	4.4	1149	4 US-09-868-552-5	Sequence 5, App11
12	46.2	4.4	1149	4 US-09-868-552-11	Sequence 11, App1
13	46.2	4.4	1149	4 US-09-868-552-13	Sequence 13, App1
14	46.2	4.4	1149	4 US-09-868-552-25	Sequence 25, App1
15	46.2	4.4	1540	4 US-09-868-552-24	Sequence 24, App1
16	46.2	4.2	962	4 US-09-868-552-45	Sequence 45, App1
17	44.6	4.2	1149	4 US-09-868-552-9	Sequence 9, App11
18	44.6	4.2	1149	4 US-09-868-552-19	Sequence 19, App1
19	44.6	4.2	1149	4 US-09-868-552-22	Sequence 22, App1
20	44.6	4.2	1530	4 US-09-868-552-21	Sequence 21, App1
21	44.6	4.2	1540	4 US-09-868-552-18	Sequence 18, App1
22	44.4	4.2	966	2 US-08-766-738-2	Sequence 2, App11
23	44.4	4.2	966	4 US-09-262-610-2	Sequence 2, App11
24	44.2	4.2	9636	1 US-08-323-170B-1	Sequence 1, App11
25	44.2	4.2	9636	3 US-08-954-441-1	Sequence 1, App11
26	43	4.0	1149	4 US-09-868-552-7	Sequence 7, App11
27	43	4.0	1149	4 US-09-868-552-16	Sequence 16, App1

28	43	4.0	1270	4 US-08-387-805-1	Sequence 1, App11
29	43	4.0	1530	4 US-09-868-552-15	Sequence 15, App1
30	43	4.0	1530	1 US-07-866-979-5	Sequence 5, App11
31	43	4.0	1533	2 US-08-466-906B-5	Sequence 5, App11
32	43	4.0	1533	3 US-08-706-281A-5	Sequence 5, App11
33	43	4.0	1533	3 US-09-201-746-5	Sequence 5, App11
34	43	4.0	1533	4 US-09-097-231-5	Sequence 5, App11
35	43	4.0	1533	4 US-09-353-099-5	Sequence 5, App11
36	43	4.0	1533	4 US-09-016-434-1343	Sequence 1343, App11
37	41.4	3.9	1235	1 US-08-696-770-1	Sequence 1, App11
38	41.4	3.9	1235	2 US-09-015-557-1	Sequence 1, App11
39	41.2	3.9	2550	4 US-08-961-527-364	Sequence 364, App1
40	40.2	3.8	1080	1 US-08-671-525B-5	Sequence 5, App11
41	40.2	3.8	1080	1 US-08-672-109B-5	Sequence 5, App11
42	40.2	3.8	1080	1 US-08-842-045-5	Sequence 5, App11
43	40.2	3.8	1080	2 US-08-842-238-5	Sequence 5, App11
44	40.2	3.8	1080	3 US-08-629-335B-5	Sequence 5, App11
45	40.2	3.8	1080	4 US-09-709-066-3	Sequence 3, App11

ALIGNMENTS

RESULT 1
US-09-627-465B-1/c
Sequence 1, Application US/09627465B
Patent No. 6737519
GENERAL INFORMATION:
APPLICANT: KEITH, TIM
APPLICANT: LITTLE, RANDALL D.
APPLICANT: BERDEWEGH, PAUL VAN
APPLICANT: DUPUIS, JOSEF
APPLICANT: DEL MASTRO, RICHARD L.
APPLICANT: SIMON, JASON
APPLICANT: ALLEN, KRISTINA
APPLICANT: PANDIT, SUNITA
TITLE OF INVENTION: NOVEL HUMAN GENES RELATING TO RESPIRATORY DISEASES AND OBESITY
FILE REFERENCE: 2976-4037
CURRENT APPLICATION NUMBER: US/09/627,465B
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/211,749
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: 60/146,336
PRIOR FILING DATE: 1999-07-30
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver 2.1
SEQ ID NO 1
LENGTH: 10304
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (267)
OTHER INFORMATION: a, t, c or g
US-09-627-465B-1
Query Match
Best Local Similarity 48.1%; Score 63.2; DB 4; Length 10304;
Pred. No. 1.8e-07;
Matches 179; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
QY 513 TGAAGACTCAGCAGACCTCTGTGATCATCGCTCCATCGTGGATCGACTGCTTACCGT 572
DB 10294 TGAAGACTCAGCAGACCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 10235
QY 573 CTACCTGTGACCTGCTCATCTTCTTCACTTGAATCAATCATTTGTAGACGCTCAG 632
DB 10234 CATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATG 10175
QY 633 GAGGAAGCAATTTTCCTCGTGGCTAGTCTCAGGGAAGACCAACCGCATCTTGT 692
DB 10174 CATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 10115
QY 693 CACCATTTACCTCATCTTTGGCAACACTTGGGCCCCCGCATCATCATCATTTTACCA 752

Db 10114 CAGCATCACCATTACATATCATCAATCAATCCGTACCAATCATCATCCGTACCAT 10055
QY 753 CCTCTATGGGGGCCCATCCAGAACCGCTGGCTGTGACATCATGTCCGACATGGCAA 812
Db 10054 CACCATCAGCATCATCATCATCATGACGACATCAACCCGTACCATCATCATCATAT 9995
QY 813 CAGCTAGCCCTTCTGAAACAGACGATCACTTCTCTCTACTGCTTCACTGACGACG 872
Db 9994 CACCATCAGCATCATCATCCGTACCATCATATCAATCAATCAACCATCATCATCAT 9935
QY 873 GTTCGCAACCAT 884
Db 9934 CATCAGCATCAT 9923

RESULT 2
US-08-233-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29pt-F18
US-08-232-463-14

Query Match 5.6%; Score 59.6; DB 1; Length 7218;
Best Local Similarity 4.3%; Pred. No. 1.6e-06;
Matches 17; Conservative 225; Mismatches 154; Indels 0; Gaps 0;
QY 461 TCACCTGCTTCTTACGACGATCCCTATTACTGTGTGCGCAACATCTGACGAGACT 520
Db 1069 YY 1128

QY 521 ACATCAGACCTCTGTGATCAGCTTCATCTGATCCAGTTCACGCTTACCTGTG 580
Db 1129 YY 1188
QY 581 TGCCCTGCTCATTCTTCTCATCTGTGAATCAATGATGTGACAGCTCAGAGAAAGA 640
Db 1189 YY 1248
QY 641 GCATTTTGTCTCCGTGGCTACTCCACGGGAGACGCGCATCTTGTTCACCATTA 700
Db 1249 YY 1308
QY 701 CCTCATCTTGTGACAGCTTGGGCCCGCATATATATATCTTTACACCTTATG 760
Db 1309 YY 1368
QY 761 GGGCGCCATCCAGAACCGCTGGCTGTGACATCATGTCCGATTCGCAACATGCTAG 820
Db 1369 YY 1428
QY 821 CCTCTGTGACAGCGCATCACTTCTCTCTACT 856
Db 1429 YYYYYYGAACAATTTCTTATCTCTTAACTACT 1464

RESULT 3
US-09-433-699-3/c
Sequence 3, Application US/09433699B
Patent No. 6165786
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Lex M. Cowsett
TITLE OF INVENTION: ANTISENSE MODULATION OF NUCLEOLIN EXPRESSION
FILE REFERENCE: RTS-0109
CURRENT APPLICATION NUMBER: US/09/433,699B
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 2518
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (112) .. (2235)
US-09-433-699-3

Query Match 4.4%; Score 46.8; DB 3; Length 2518;
Best Local Similarity 46.6%; Pred. No. 0.004;
Matches 150; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
QY 543 CGTCTCATCTGATTCATCTGTTCAACGCTCTACTGTGCTGCTCTCATCTTCTTAT 602
Db 880 CATCATCTTCATCATCTTCTGCTGTGTGTCTCTCTCTCTCATCATCTTCTTCTAT 821
QY 603 CTGGAATCAATCATGTGTGTAACAAGCTCAGAGAGAAATTTGCTCCGGGTGA 662
Db 820 CTTCATCTCAGACGAGTCTTGCTTCAACAGAAACATTTGACACTTCTTCTT 761
QY 663 CTCACGGGAGAGACACCGCATTTGTTCACATTAATCTCAATCTTTCACACTTTG 722
Db 760 TGGCTGGTGTATCTTCATAGCTTCTTCTTCAGAGTCACTTCTTCATGCTATGCTAT 701
QY 723 GGGCCCCGACATCATGATTTCTTACCACTCTATAGGGGGCGCCATCCAGAACGCTG 782
Db 700 CCTCATCATCTTCTGATCTCATCTCATCTCTATCTCTGAGGAGGAGGAGAGCTGCTG 641
QY 783 GGTGTGACATCATGTCCGACATTCGCAACATGCTAATCTTCTGAAACAGGCATCAA 842
Db 640 CTTCATGCTGCTGCTTCAATTCATCTTCATCTCATCTGCTCTCTCTCTCTCTCT 581
QY 843 CTTCCTCTTACTGCTTCAATC 864
Db 580 CATCTCTCATCTCATCATC 559

RESULT 4

US-08-671-525B-1
Sequence 1, Application US/08671525B
Patent No. 5703220
GENERAL INFORMATION:
APPLICANT: Yamada, Tadataka
APPLICANT: Gantz, Ira
TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: US
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671,525B
FILING DATE: June 27, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Dean F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-000853DVB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810)641-1600
TELEFAX: (810)641-0270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 951 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 1..951
FEATURE:
NAME/KEY: conflict
LOCATION: order(269..270, 488, 490..491)
OTHER INFORMATION: /note= "Differs from sequence
OTHER INFORMATION: published by Chhajlani and Wilberg in
OTHER INFORMATION: five nucleotide and three amino acid
OTHER INFORMATION: positions."
FEATURE:
NAME/KEY: conflict
LOCATION: order(485, 488)
OTHER INFORMATION: /note= "Differs from sequence
OTHER INFORMATION: published by Mountjoy, Robbins,
OTHER INFORMATION: Mortrud and Cone in 2 nucleotide and 2
OTHER INFORMATION: amino acid positions."
US-08-671-525B-1

Query Match 4.4%; Score 46.2; DB 1; Length 951;
Best Local Similarity 51.7%; Pred. No. 0.0034;

Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGGTCCTCCGACAGATCTTAAAGTCTGGAATTCTATCCATCCACCTC 344
DB 333 GGCGCTCAGGTCCTCCGACAGATCTTAAAGTCTGGAATTCTATCCATCCACCTC 392
QY 345 CATATGATTAATCTGTAACGTTAACAATTGACAGATTAATGCTGTGCGACACCGCTCA 404

RESULT 5

US-08-672-109B-1
Sequence 1, Application US/08672109B
Patent No. 5710265
GENERAL INFORMATION:
APPLICANT: Yamada, Tadataka
APPLICANT: Gantz, Ira
TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: US
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/672,109B
FILING DATE: June 27, 1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Dean F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-000853DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810)641-1600
TELEFAX: (810)641-0270
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 951 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 1..951
FEATURE:
NAME/KEY: conflict
LOCATION: order(269..270, 488, 490..491)
OTHER INFORMATION: /note= "Differs from sequence
OTHER INFORMATION: published by Chhajlani and Wilberg in
OTHER INFORMATION: five nucleotide and three amino acid
OTHER INFORMATION: positions."
FEATURE:
NAME/KEY: conflict
LOCATION: order(485, 488)
OTHER INFORMATION: /note= "Differs from sequence
OTHER INFORMATION: published by Mountjoy, Robbins,
OTHER INFORMATION: Mortrud and Cone in 2 nucleotide and 2
OTHER INFORMATION: amino acid positions."
US-08-672-109B-1

DB 393 CCTCTCTCTCTGAGGCGCATGCGCCGTGAGACCGCTACATCTTCATCTTACGCACTGCG 452
QY 405 GTACACACAGGTCCTCATACCCAGCCGCAACCGGAAAGTCTGTAAGTTTACATCAC 464
DB 453 CTACACACAGATGTAACCTGCGCGGCGGCAAGCCGTTGCGGCATCTGGGTGGC 512
QY 465 CTGCTTCCTGACACAGATCCCT 487
DB 513 CAGTGTGCTTTCAGACAGCTCT 535


```

; TYPE: DNA
; ORGANISM: Homo sapien
US-09-868-552-1

Query Match
Best Local Similarity 4.4%; Score 46.2; DB 4; Length 1149;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGGTCCCGGACAGATCATAGAGTCTGGAATTTCTCATCTCCACACCTC 344
Db 333 GGCTCTCAGAGGTGACATGTCATGATGATACCTGCAGCTCCATGCTGTCAG 392
QY 345 CATATGATTACTGTACCTGTAACCATTTGACAGGTATATGCTGTGACACCGGCTCA 404
Db 393 CCTCTGCTTCTGGGGGCGCATCGCGTGACCGGTATATCTTCATCTTCAACGCACTGCG 452
QY 405 GTACCAACAGGTCTCATATCCAGCCGACCCGAAAGTCAATTGTAAGTTTACATCAC 464
Db 453 TTACCAACAGCATGTGACCTTGCCTGGCGGCGGCAAGCGGCTTGCGGCTGGG 512
QY 465 CTGCTTCTGACAGCATCCCT 487
Db 513 CAGTGTCTTTCAGCAGCTCT 535

RESULT 10
US-09-868-552-3
; Sequence 3, Application US/09868552
; Patent No. 6693184
; GENERAL INFORMATION:
; APPLICANT: Howard, Andrew D.
; APPLICANT: Macneil, Douglas J.
; TITLE OF INVENTION: DNA MOLECULES ENCODING SPICE VARIANTS
; FILE REFERENCE: 20367P
; CURRENT APPLICATION NUMBER: US/09/868,552
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: PCT/US99/29963
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,401
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-868-552-3

Query Match
Best Local Similarity 4.4%; Score 46.2; DB 4; Length 1149;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGGTCCCGGACAGATCATAGAGTCTGGAATTTCTCATCTCCACACCTC 344
Db 333 GGCTCTCAGAGGTGACATGTCATGATGATACCTGCAGCTCCATGCTGTCAG 392
QY 345 CATATGATTACTGTACCTGTAACCATTTGACAGGTATATGCTGTGACACCGGCTCA 404
Db 393 CCTCTGCTTCTGGGGGCGCATCGCGTGACCGGTATATCTTCATCTTCAACGCACTGCG 452
QY 405 GTACCAACAGGTCTCATATCCAGCCGACCCGAAAGTCAATTGTAAGTTTACATCAC 464
Db 453 TTACCAACAGCATGTGACCTTGCCTGGCGGCGGCAAGCGGCTTGCGGCTGGG 512
QY 465 CTGCTTCTGACAGCATCCCT 487
Db 513 CAGTGTCTTTCAGCAGCTCT 535

RESULT 11
US-09-868-552-5
```

```

; Sequence 5, Application US/09868552
; Patent No. 6693184
; GENERAL INFORMATION:
; APPLICANT: Howard, Andrew D.
; APPLICANT: Macneil, Douglas J.
; TITLE OF INVENTION: DNA MOLECULES ENCODING SPICE VARIANTS
; FILE REFERENCE: 20367P
; CURRENT APPLICATION NUMBER: US/09/868,552
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: PCT/US99/29963
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,401
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-868-552-5

Query Match
Best Local Similarity 4.4%; Score 46.2; DB 4; Length 1149;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 285 GATGCTCAGGTCCCGGACAGATCATAGAGTCTGGAATTTCTCATCTCCACACCTC 344
Db 333 GGCTCTCAGAGGTGACATGTCATGATGATACCTGCAGCTCCATGCTGTCAG 392
QY 345 CATATGATTACTGTACCTGTAACCATTTGACAGGTATATGCTGTGACACCGGCTCA 404
Db 393 CCTCTGCTTCTGGGGGCGCATCGCGTGACCGGTATATCTTCATCTTCAACGCACTGCG 452
QY 405 GTACCAACAGGTCTCATATCCAGCCGACCCGAAAGTCAATTGTAAGTTTACATCAC 464
Db 453 TTACCAACAGCATGTGACCTTGCCTGGCGGCGGCAAGCGGCTTGCGGCTGGG 512
QY 465 CTGCTTCTGACAGCATCCCT 487
Db 513 CAGTGTCTTTCAGCAGCTCT 535

RESULT 12
US-09-868-552-11
; Sequence 11, Application US/09868552
; Patent No. 6693184
; GENERAL INFORMATION:
; APPLICANT: Howard, Andrew D.
; APPLICANT: Macneil, Douglas J.
; TITLE OF INVENTION: DNA MOLECULES ENCODING SPICE VARIANTS
; FILE REFERENCE: 20367P
; CURRENT APPLICATION NUMBER: US/09/868,552
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: PCT/US99/29963
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,401
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-868-552-11

Query Match
Best Local Similarity 4.4%; Score 46.2; DB 4; Length 1149;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
```

RESULT 14
 US-09-868-552-25
 ; Sequence 25, Application US/09868552
 ; Patent No. 6693184
 ; GENERAL INFORMATION:
 ; APPLICANT: Howard, Andrew D.
 ; APPLICANT: Macneil, Douglas J.
 ; APPLICANT: Van Der Ploeg, Leonardus H. T.
 ; TITLE OF INVENTION: DNA MOLECULES ENCODING SPLICING VARIANTS
 ; TITLE OF INVENTION: OF THE HUMAN MELANOCORTIN 1 RECEPTOR PROTEIN

Query Match 4.4%; Score 46.2; DB 4; Length 1540;
Best Local Similarity 51.7%; Pred. No. 0.0045;
Matches 105; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY	345	CATATGGATTACATGTAACCGTTAAACATTATACAGATATCGCTGTCTGCCACCCGCTCAA	404
Db	393	CCCTTGCTTCTCGGCGCGCGCTCGCGCTGGACCGCTACATCTTCATCTTCTACGACATCGCG	452
QY	405	GTACACACAGCGTTCATATACCGACCCCGCACCCCGGAAGTCATTGTGAATGTTTACATCAC	464
Db	453	CTACACACAGATGTATACCTTGCCTGGCGGAGCGCGGACCGCTTGCAGGCACTGGGGTGC	512
QY	465	CTGCTTCTCGACACGATCCCTCT	487
Db	513	CAGTGTGCTTTCAGACACCTCT	535

```
Search completed: December 15, 2004, 02:56:17
Job time : 90.3365 secs
```


PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 1062
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Artificial Sequence
OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-15

Query Match 100.0%; Score 1062; DB 9; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2e-310;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACACGACGCGCCACCTGCGACGCAAGCTGCGTCTTGGTGGTCCCGCGC 60
DB 1 ATGAGACACGACGCGCCACCTGCGACGCAAGCTGCGTCTTGGTGGTCCCGCGC 60
QY 61 TCGGCTGCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
DB 61 TCGGCTGCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
QY 121 TTACGACCAAAATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAA 180
DB 121 TTACGACCAAAATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAA 180
QY 181 TCTCTCCACAACTATCTTGGGCACTGCTGCGCGACACTTGGTCTCTTTTCATA 240
DB 181 TCTCTCCACAACTATCTTGGGCACTGCTGCGCGACACTTGGTCTCTTTTCATA 240
QY 241 GTGTTGTGAGCTTCTGTTGAAGATTCATCTTGAAGATTCATCTTGAAGATTCAT 300
DB 241 GTGTTGTGAGCTTCTGTTGAAGATTCATCTTGAAGATTCATCTTGAAGATTCAT 300
QY 301 GACAAATCATAGAGTGTGGAATTCATCTTGAAGATTCATCTTGAAGATTCAT 360
DB 301 GACAAATCATAGAGTGTGGAATTCATCTTGAAGATTCATCTTGAAGATTCAT 360
QY 361 CGCTTACCAATGACAGTATTCGCTGCTGCGACCGCTCAAGTACACACGCTTCA 420
DB 361 CGCTTACCAATGACAGTATTCGCTGCTGCGACCGCTCAAGTACACACGCTTCA 420
QY 421 TACCCAGCCGCAACCGGAAAGTCAATGTAAGTGTGAATCACTGCTGCGACG 480
DB 421 TACCCAGCCGCAACCGGAAAGTCAATGTAAGTGTGAATCACTGCTGCGACG 480
QY 481 ATCCCTATTAATGTTGGTGGCCCAATCTGGAATGAAGTATGATGATGATGAT 540
DB 481 ATCCCTATTAATGTTGGTGGCCCAATCTGGAATGAAGTATGATGATGATGAT 540
QY 541 CAGCTCTCATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 600
DB 541 CAGCTCTCATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 600
QY 601 ATCTTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ATCTTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 TACTCAGGAGGAGACACGCGCATCTGTTGACATTAATCTGATCTTGGCACTT 720
DB 661 TACTCAGGAGGAGACACGCGCATCTGTTGACATTAATCTGATCTTGGCACTT 720
QY 721 TGGGCGCGCGCATCATGATCTTGAACCTCTTGAAGGCGCCATCAGAACCGC 780
DB 721 TGGGCGCGCGCATCATGATCTTGAACCTCTTGAAGGCGCCATCAGAACCGC 780
QY 781 TGGGCGCGCGCATCATGATCTTGAACCTCTTGAAGGCGCCATCAGAACCGC 840
DB 781 TGGGCGCGCGCATCATGATCTTGAACCTCTTGAAGGCGCCATCAGAACCGC 840

QY 841 AACTCTCTCTACTGCTTCAATCAGAACGCGGTTCCGACCATGCGACCGCCACGCTC 900
DB 841 AACTCTCTCTACTGCTTCAATCAGAACGCGGTTCCGACCATGCGACCGCCACGCTC 900
QY 901 AAGGCTTTCTTCAAGTCCGAGAACCTGTGACAGTTCTTACCAATCATTAATTTCC 960
DB 901 AAGGCTTTCTTCAAGTCCGAGAACCTGTGACAGTTCTTACCAATCATTAATTTCC 960
QY 961 ATTAACAGTACGCGCTGATCTGCGCGGCAAACTACACTGATCAAGATGCTGTAC 1020
DB 961 ATTAACAGTACGCGCTGATCTGCGCGGCAAACTACACTGATCAAGATGCTGTAC 1020
QY 1021 CAGTATGACAAATAATGAAAACTTAAAGATCCCGGTGA 1062
DB 1021 CAGTATGACAAATAATGAAAACTTAAAGATCCCGGTGA 1062

RESULT 2

US-09-995-225-15
Sequence 15, Application US/09995225
GENERAL INFORMATION:
APPLICANT: Chem. Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huang T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Pridle, Cameron
TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
FILE REFERENCE: Apen-0308
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 1062
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20030139588A9el Sequence
US-09-995-225-15

Query Match 100.0%; Score 1062; DB 10; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2e-310;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACACGACGCGCCACCTGCGACGCAAGCTGCGTCTTGGTGGTCCCGCGC 60

Db 1 ATGGAGACACGACGCCACCTCGAGCCACAGCTCGCTGTTGTGTGTCCTCCCGC 60
QY 61 TCGGCTGCGGCTTGGGTTTGTGCCCCGTCTACTACAGCTCTTGTGCTCTCGT 120
Db 61 TCGGCTGCGGCTTGGGTTTGTGCCCCGTCTACTACAGCTCTTGTGCTCTCGT 120
QY 121 TTACGACGAAATCTTGAAGATGATCTCTCCAGCTGTGGCAAGACAGAG 180
Db 121 TTACGACGAAATCTTGAAGATGATCTCTCCAGCTGTGGCAAGACAGAG 180
QY 181 TCCTCTACAACTATCTTGGCACTCGCTGCGCAACATCTTGGCTCTTTTCA 240
Db 181 TCCTCTACAACTATCTTGGCACTCGCTGCGCAACATCTTGGCTCTTTTCA 240
QY 241 GTGTTGTGAGCTTCTGTGGAAGATTTTCACTTGAACATGACAGATGCTCAGTCCC 300
Db 241 GTGTTGTGAGCTTCTGTGGAAGATTTTCACTTGAACATGACAGATGCTCAGTCCC 300
QY 301 GACAAATCATTAAGTGTGGAATTCATTCATCCATCCACCTCCCAATGATTAATCTGTA 360
Db 301 GACAAATCATTAAGTGTGGAATTCATTCATCCATCCACCTCCCAATGATTAATCTGTA 360
QY 361 CCGTTAACCATTAAGATGATATCGCTGTGCAACCGCTCAAGTACCAAGGTCTCA 420
Db 361 CCGTTAACCATTAAGATGATATCGCTGTGCAACCGCTCAAGTACCAAGGTCTCA 420
QY 421 TACCCAGCCCGCAACCGGAAAGTCAATTGTAAGTGTTAACATCACTGTCTCTGACACG 480
Db 421 TACCCAGCCCGCAACCGGAAAGTCAATTGTAAGTGTTAACATCACTGTCTCTGACACG 480
QY 481 ATCCCTATTAATCTGATGAGCCCAACATCTGCACTGAAGACATCAACAGCACTCTGAT 540
Db 481 ATCCCTATTAATCTGATGAGCCCAACATCTGCACTGAAGACATCAACAGCACTCTGAT 540
QY 541 CAGCTCTCATCTGATATCCACTGCTTCAACGCTCTACCTGCTGCTCTGCTCTTCTTC 600
Db 541 CAGCTCTCATCTGATATCCACTGCTTCAACGCTCTACCTGCTGCTCTGCTCTTCTTC 600
QY 601 ATCTTGAATCAATCAATTTGTGTAACAAGCTCAGAGAGAGCAATTTTGTCTCCGTGAC 660
Db 601 ATCTTGAATCAATCAATTTGTGTAACAAGCTCAGAGAGAGCAATTTTGTCTCCGTGAC 660
QY 661 TACTCCAGCGGGAAGACACCGGCAATCTTGTCAACCTTACCTTCTTGTGCAACTT 720
Db 661 TACTCCAGCGGGAAGACACCGGCAATCTTGTCAACCTTACCTTCTTGTGCAACTT 720
QY 721 TGGGCCCCCGCATCATGATTTCTTTACCACTCTCTATGAGGCGCCCATCAGAAACCG 780
Db 721 TGGGCCCCCGCATCATGATTTCTTTACCACTCTCTATGAGGCGCCCATCAGAAACCG 780
QY 781 TGGCTGTGATCATATGTCGACATTTGCCAAGATGCTAGCCCTTGTGAACAGACCATC 840
Db 781 TGGCTGTGATCATATGTCGACATTTGCCAAGATGCTAGCCCTTGTGAACAGACCATC 840
QY 841 AACTTTCTCTCTCATCTGCTTCAATCAGAAAGCGGTTCGCAACATGAGCGGCAAGCTC 900
Db 841 AACTTTCTCTCTCATCTGCTTCAATCAGAAAGCGGTTCGCAACATGAGCGGCAAGCTC 900
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACACCAATCAATCTTTTCC 960
Db 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACACCAATCAATCTTTTCC 960
QY 961 ATTAACAAGTAGCCCTGAGATCTGCGCGGCAAACTCACTGATCAAGATGCTGTGTAC 1020
Db 961 ATTAACAAGTAGCCCTGAGATCTGCGCGGCAAACTCACTGATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAATAATGAAAACTATTAATAAGTATCCCGGTGA 1062
Db 1021 CAGTATGACAAATAATGAAAACTATTAATAAGTATCCCGGTGA 1062

RESULT 3
US-10-296-294A-4

Sequence 4, Application US/10296294A
Publication No. US20040029224A1
GENERAL INFORMATION:
APPLICANT: TERAO, Yasuko
APPLICANT: MATSUI, Hideki
APPLICANT: SHINTANI, Yasushi
TITLE OF INVENTION: No. US20040029224A1e1 G Protein-Coupled Receptor and its DNA
FILE REFERENCE: 2734 USOP
CURRENT APPLICATION NUMBER: US/10/296,294A
CURRENT FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: PCT/JP01/04643
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: JP 2000-170446
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: JP 2000-194926
PRIOR FILING DATE: 2000-06-23
NUMBER OF SEQ ID NOS: 8
SEQ ID NO 4
LENGTH: 1062
TYPE: DNA
ORGANISM: Human
US-10-296-294A-4
Query Match 100.0%; Score 1062; DB 16; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2e-310;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGACACGACGCCACCTCGAGCCACAGCTCGCTGTTGTGTGTCCTCCCGC 60
Db 1 ATGGAGACACGACGCCACCTCGAGCCACAGCTCGCTGTTGTGTGTCCTCCCGC 60
QY 61 TCGGCTGCGGCTTGGGTTTGTGCCCCGTCTACTACAGCTCTTGTGCTCTCGT 120
Db 61 TCGGCTGCGGCTTGGGTTTGTGCCCCGTCTACTACAGCTCTTGTGCTCTCGT 120
QY 121 TTACGACGAAATCTTGAAGATGATCTCTCCAGCTGTGGCAAGACAGAG 180
Db 121 TTACGACGAAATCTTGAAGATGATCTCTCCAGCTGTGGCAAGACAGAG 180
QY 181 TCCTCTACAACTATCTTGGCACTCGCTGCGCAACATCTTGGCTCTTTTCA 240
Db 181 TCCTCTACAACTATCTTGGCACTCGCTGCGCAACATCTTGGCTCTTTTCA 240
QY 241 GTGTTGTGAGCTTCTGTGGAAGATTTTCACTTGAACATGACAGATGCTCAGTCCC 300
Db 241 GTGTTGTGAGCTTCTGTGGAAGATTTTCACTTGAACATGACAGATGCTCAGTCCC 300
QY 301 GACAAATCATTAAGTGTGGAATTCATTCATCCATCCACCTCCCAATGATTAATCTGTA 360
Db 301 GACAAATCATTAAGTGTGGAATTCATTCATCCATCCACCTCCCAATGATTAATCTGTA 360
QY 361 CCGTTAACCATTAAGATGATATCGCTGTGCAACCGCTCAAGTACCAAGGTCTCA 420
Db 361 CCGTTAACCATTAAGATGATATCGCTGTGCAACCGCTCAAGTACCAAGGTCTCA 420
QY 421 TACCCAGCCCGCAACCGGAAAGTCAATTGTAAGTGTTAACATCACTGTCTCTGACACG 480
Db 421 TACCCAGCCCGCAACCGGAAAGTCAATTGTAAGTGTTAACATCACTGTCTCTGACACG 480
QY 481 ATCCCTATTAATCTGATGAGCCCAACATCTGCACTGAAGACATCAACAGCACTCTGAT 540
Db 481 ATCCCTATTAATCTGATGAGCCCAACATCTGCACTGAAGACATCAACAGCACTCTGAT 540
QY 541 CAGCTCTCATCTGATATCCACTGCTTCAACGCTCTACCTGCTGCTCTGCTCTTCTTC 600
Db 541 CAGCTCTCATCTGATATCCACTGCTTCAACGCTCTACCTGCTGCTCTGCTCTTCTTC 600
QY 601 ATCTTGAATCAATCAATTTGTGTAACAAGCTCAGAGAGAGCAATTTTGTCTCCGTGAC 660
Db 601 ATCTTGAATCAATCAATTTGTGTAACAAGCTCAGAGAGAGCAATTTTGTCTCCGTGAC 660
QY 661 TACTCCAGCGGGAAGACACCGGCAATCTTGTCAACCTTACCTTCTTGTGCAACTT 720
Db 661 TACTCCAGCGGGAAGACACCGGCAATCTTGTCAACCTTACCTTCTTGTGCAACTT 720

Db 661 TACTCCACGGGGAAGAACCGCCATCTTTTCACATTACCTCCATCTTTGCCACACTT 720
Qy 721 TGGGCCCCCGGCAATCATGATGATTTCTTTCACACTTATGAGGCGCCCATCCAGAACCGC 780
Db 721 TGGGCCCCCGGCAATCATGATGATTTCTTTCACACTTATGAGGCGCCCATCCAGAACCGC 780
Qy 781 TGGCTGGTGCATCATGATGATTCGACATTCGCAACATGCTTAAGCCCTTGAACACAGCCATC 840
Db 781 TGGCTGGTGCATCATGATGATTCGACATTCGCAACATGCTTAAGCCCTTGAACACAGCCATC 840
Qy 841 AACTTCTTCTCTACTGCTTTCATACGAAAGCGGTTCCGACCATGAGCGCCACAGCTC 900
Db 841 AACTTCTTCTCTACTGCTTTCATACGAAAGCGGTTCCGACCATGAGCGCCACAGCTC 900
Qy 901 AAGGCTTTCTTCAAGTGCAGAACGCTGATGCTTCAAGCTTCAACCATCATTAATCTTTTCC 960
Db 901 AAGGCTTTCTTCAAGTGCAGAACGCTGATGCTTCAAGCTTCAACCATCATTAATCTTTTCC 960
Qy 961 ATAAACAAGTAAAGCCCTGATCTCGCGGCAAACTCACATGCTCATCAAGATGCTGTATAC 1020
Db 961 ATAAACAAGTAAAGCCCTGATCTCGCGGCAAACTCACATGCTCATCAAGATGCTGTATAC 1020
Qy 1021 CAGTATGACAAAATGAAAACTATAAAGTATCCCGTGA 1062
Db 1021 CAGTATGACAAAATGAAAACTATAAAGTATCCCGTGA 1062

RESULT 4

US-10-779-104-1
; Sequence 1, Application US/10779104
; Publication No. US20040161799A1
; GENERAL INFORMATION:
; APPLICANT: Susan Croll-Kalish
; APPLICANT: Andrew J. Murphy
; TITLE OF INVENTION: KOR3like-Proteins and Methods of Modulating KOR3L-Mediated Activity
; FILE REFERENCE: REG 1000A
; CURRENT APPLICATION NUMBER: US/10/779,104
; PRIOR FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: 60/447,447
; PRIOR FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/495,577
; PRIOR FILING DATE: 2003-08-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-779-104-1

Query Match 100.0%; Score 1062; DB 17; Length 2117;
Best Local Similarity 100.0%; Pred. No. 2.8e-310;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAGCAACAGCGACGCCCACTCGAGCAACAGCTCGCTGTCTTGTGGTCCCGGCGC 60
Db 280 ATGGAGCAACAGCGACGCCCACTCGAGCAACAGCTCGCTGTCTTGTGGTCCCGGCGC 339
Qy 61 TCGGCTGGGCGCTTGGGTTTGGTGGCGAGGCTTCACTCAACCTCTTGTGCTGCTGGT 120
Db 340 TCGGCTGGGCGCTTGGGTTTGGTGGCGAGGCTTCACTCAACCTCTTGTGCTGCTGGT 399
Qy 121 TTACCAAGCAATATTTTACAGTATCATCTCTCCAGCTGGTGGCAAGAAAGCAGAG 180
Db 400 TTACCAAGCAATATTTTACAGTATCATCTCTCCAGCTGGTGGCAAGAAAGCAGAG 459
Qy 181 TCTCTCTCAACTATCTTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 460 TCTCTCTCAACTATCTTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
Qy 241 GTGTTTGTGACTTCTGTTGGAAGATTTCATTTGAACATGCAAGATGCTGCTGCTGCTGCT 300
Db 520 GTGTTTGTGACTTCTGTTGGAAGATTTCATTTGAACATGCAAGATGCTGCTGCTGCTGCTGCT 579

Qy 301 GACAAAGTCAATAGAAGTGTGGAATTTCTATCATCAACCTCCATATGATTAAGTACTGA 360
Db 580 GACAAAGTCAATAGAAGTGTGGAATTTCTATCATCAACCTCCATATGATTAAGTACTGA 639
Qy 361 CCGTTAAACATTGACAGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 640 CCGTTAAACATTGACAGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699
Qy 421 TACCCAGCCCGCAACCCGGAAGTCAATGTAAGTTTATCATCACTGCTTCTGACAGC 480
Db 700 TACCCAGCCCGCAACCCGGAAGTCAATGTAAGTTTATCATCACTGCTTCTGACAGC 759
Qy 481 ATCCCTATTTATGAGGCGCCCAACATCTGAGCTGAAGCTACATCAAGCACTTGTGAT 540
Db 760 ATCCCTATTTATGAGTGGCCCAACATCTGAGCTGAAGCTACATCAAGCACTTGTGAT 819
Qy 541 CAGGCTCTCATCTGATCACTGCTTCAACGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 820 CAGGCTCTCATCTGATCACTGCTTCAACGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 879
Qy 601 ATCTTGAATCAATGATGTTATCAAGTCAAGGAGAAAGCAATTTTGTCTCGAGGC 660
Db 880 ATCTTGAATCAATGATGTTATCAAGTCAAGGAGAAAGCAATTTTGTCTCGAGGC 939
Qy 661 TACTCCACGGGGAAGAACCGCCATCTTGTTCACATTAACCTTCATCTTTGCAACACTT 720
Db 940 TACTCCACGGGGAAGAACCGCCATCTTGTTCACATTAACCTTCATCTTTGCAACACTT 999
Qy 721 TGGGCCCCCGGCAATCATGATGATTTCTTTCACACTTATGAGGCGCCCATCCAGAACCGC 780
Db 1000 TGGGCCCCCGGCAATCATGATGATTTCTTTCACACTTATGAGGCGCCCATCCAGAACCGC 1059
Qy 781 TGGCTGGTGCATCATGATGATTCGACATTCGCAACATGCTTAAGCCCTTGAACACAGCCATC 840
Db 1060 TGGCTGGTGCATCATGATGATTCGACATTCGCAACATGCTTAAGCCCTTGAACACAGCCATC 1119
Qy 841 AACTTCTTCTCTACTGCTTTCATACGAAAGCGGTTCCGACCATGAGCGCCACAGCTC 900
Db 1120 AACTTCTTCTCTACTGCTTTCATACGAAAGCGGTTCCGACCATGAGCGCCACAGCTC 1179
Qy 901 AAGGCTTTCTTCAAGTGCAGAACGCTGATGCTTCAAGCTTCAACCATCATTAATCTTTTCC 960
Db 1180 AAGGCTTTCTTCAAGTGCAGAACGCTGATGCTTCAAGCTTCAACCATCATTAATCTTTTCC 1239
Qy 961 ATAAACAAGTAAAGCCCTGATCTCGCGGCAAACTCACATGCTCATCAAGATGCTGTATAC 1020
Db 1240 ATAAACAAGTAAAGCCCTGATCTCGCGGCAAACTCACATGCTCATCAAGATGCTGTATAC 1299
Qy 1021 CAGTATGACAAAATGAAAACTATAAAGTATCCCGTGA 1062
Db 1300 CAGTATGACAAAATGAAAACTATAAAGTATCCCGTGA 1341

RESULT 5

US-10-296-294A-3
; Sequence 3, Application US/10296294A
; Publication No. US20040029224A1
; GENERAL INFORMATION:
; APPLICANT: TERAO, Yasuko
; APPLICANT: MATSUI, Hideki
; TITLE OF INVENTION: SHINTANI, Yasuhide
; FILE REFERENCE: 2734 USOP
; CURRENT APPLICATION NUMBER: US/10/296,294A
; PRIOR FILING DATE: 2001-11-21, 294A
; PRIOR APPLICATION NUMBER: PCT/JP01/04643
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: JP 2000-170446
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: JP 2000-194926
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 8

SEQ ID NO 3
LENGTH: 1062
TYPE: DNA
ORGANISM: Human
US-10-296-294A-3

Query Match 99.8%; Score 1060.4; DB 16; Length 1062;
Best Local Similarity 99.9%; Pred. No. 6e-310; Indels 0; Gaps 0;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGACACACGACGCGCCACCTCGAGCCAGACGCTCGTCTTGTTGTCCTCCCGCGC 60
DB 1 ATGAGACACACGACGCGCCACCTCGAGCCAGACGCTCGTCTTGTTGTCCTCCCGCGC 60
QY 61 TCGGCTGCGGCTTGGGTTTGTGCGCCGCTGCTACTACAGCCTCTTGCTGCTCGGT 120
DB 61 TCGGCTGCGGCTTGGGTTTGTGCGCCGCTGCTACTACAGCCTCTTGCTGCTCGGT 120
QY 121 TTACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAGAG 180
DB 121 TTACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAGAGAG 180
QY 181 TCCTCTACAACTATCTCTTGGACATCGCTGCGGACATCTTGGTCTCTTTTCATA 240
DB 181 TCCTCTACAACTATCTCTTGGACATCGCTGCGGACATCTTGGTCTCTTTTCATA 240
QY 241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGCAATGCTCAGTCCC 300
DB 241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGCAATGCTCAGTCCC 300
QY 301 GACAGATCATAGAGGTGGAATTCATCATGACACACCTCCATATGATGATCTGTA 360
DB 301 GACAGATCATAGAGGTGGAATTCATCATGACACACCTCCATATGATGATCTGTA 360
QY 361 CCGTTAACATGACAGATATATGCTGTCTGACCCGCTCAAGTACACAGCTCTCA 420
DB 361 CCGTTAACATGACAGATATATGCTGTCTGACCCGCTCAAGTACACAGCTCTCA 420
QY 421 TACCCAGCCGCGACCCGGAAGTTCATGTTTAACTATCACTGCTTCTGACGAC 480
DB 421 TACCCAGCCGCGACCCGGAAGTTCATGTTTAACTATCACTGCTTCTGACGAC 480
QY 481 ATCCCTATATAGTGTGCGCCACATCTGACATGAGACATCACTGATGAT 540
DB 481 ATCCCTATATAGTGTGCGCCACATCTGACATGAGACATCACTGATGAT 540
QY 541 CAGTCTCATCTGATTCACCTGCTTACCGTCTACCTGCTGCTGCTCATCTTCT 600
DB 541 CAGTCTCATCTGATTCACCTGCTTACCGTCTACCTGCTGCTGCTCATCTTCT 600
QY 601 ATTTGAATCATCTTGTGTGACAGCTGACAGAGAGAGCAATTTTGTCTCGTGC 660
DB 601 ATTTGAATCATCTTGTGTGACAGCTGACAGAGAGAGCAATTTTGTCTCGTGC 660
QY 661 TACTCAGCGGGAAGACACCGCCATCTTGTGACCACTTACCTTGTGACCACT 720
DB 661 TACTCAGCGGGAAGACACCGCCATCTTGTGACCACTTACCTTGTGACCACT 720
QY 721 TGGGCCCCCGCATCATGATTTCTTACCACTGATGAGGCGCCATTCAGAAC 780
DB 721 TGGGCCCCCGCATCATGATTTCTTACCACTGATGAGGCGCCATTCAGAAC 780
QY 781 TGGCTGTGACATCATGTCGACATTTGCAATGCTGACCTTCTGAAACAGCAG 840
DB 781 TGGCTGTGACATCATGTCGACATTTGCAATGCTGACCTTCTGAAACAGCAG 840
QY 841 AACTTCTCTCTCACTGCTTCACTGACAGCGGTTCCGACATGAGAGCGGCGCAG 900
DB 841 AACTTCTCTCTCACTGCTTCACTGACAGCGGTTCCGACATGAGAGCGGCGCAG 900
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTAAGTTTCAACCAATCAATCTTTCC 960
DB 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTAAGTTTCAACCAATCAATCTTTCC 960
```

```
QY 961 ATTAACAAGTACGCGCTGATCTTCCGCGCAAACTCACTGATCAAGATGCTGTAC 1020
DB 961 ATTAACAAGTACGCGCTGATCTTCCGCGCAAACTCACTGATCAAGATGCTGTAC 1020
QY 1021 CAGTATGACAAATAATGAATACTTAAAGATGATCCCGTGA 1062
DB 1021 CAGTATGACAAATAATGAATACTTAAAGATGATCCCGTGA 1062
```

RESULT 6

```
US-10-094-417-1
; Sequence 1, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiayang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030045685A1e1 Receptors
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent Ver. 2.1
```

```
SEQ ID NO 1
LENGTH: 1202
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
NAME/KEY: CDS
LOCATION: (68)..(1129)
OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
US-10-094-417-1
```

Query Match 99.8%; Score 1060.4; DB 14; Length 1202;
Best Local Similarity 99.9%; Pred. No. 6.4e-310; Indels 0; Gaps 0;
Matches 1061; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGACACACGACGCGCCACCTCGAGCCAGACGCTCGTCTTGTTGTCCTCCCGCGC 60
DB 68 ATGAGACACACGACGCGCCACCTCGAGCCAGACGCTCGTCTTGTTGTCCTCCCGCGC 127
QY 61 TCGGCTGCGGCTTGGGTTTGTGCGCCGCTGCTACTACAGCCTCTTGCTGCTCGGT 120
DB 128 TCGGCTGCGGCTTGGGTTTGTGCGCCGCTGCTACTACAGCCTCTTGCTGCTCGGT 187
QY 121 TTACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAGAG 180
DB 188 TTACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAGAGAG 247
QY 181 TCCTCTACAACTATCTTGGACATCGCTGCGGACATCTTGGTCTCTTTTCATA 240
DB 248 TCCTCTACAACTATCTTGGACATCGCTGCGGACATCTTGGTCTCTTTTCATA 307
QY 241 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGCAATGCTCAGTCCC 300
DB 308 GTGTTTGTGACCTTCTGTTGGAAGATTTCATCTTGAACATGCAATGCTCAGTCCC 367
QY 301 GACAGATCATAGAGGTGGAATTCATGATCATCACTGCACTCAATATGATGATCTGTA 360
DB 368 GACAGATCATAGAGGTGGAATTCATGATCATCACTGCACTCAATATGATGATCTGTA 427
QY 361 CCGTTAACATGACAGGATATGCTGTGCGACCCGCTCAAGTACACACGCTTCA 420
DB 428 CCGTTAACATGACAGGATATGCTGTGCGACCCGCTCAAGTACACACGCTTCA 487
```


Qy	781	TGGCTGGGACATCATATGTCGCAATTTGGCAACATGTCAGCCCTTGTGAACAAGGCATC	84.0
Db	782	TGGCTGGGACATCATATGTCGCAATTTGGCAACATGTCAGCCCTTGTGAACAAGGCATC	84.1
Qy	841	AACTTCTTCTCTACTGCTTCACTCAGCAAGCGGTTCCGACCCATGGCAGCGGCACGCTC	90.0
Db	842	AACTTCTTCTCTACTGCTTCACTCAGCAAGCGGTTCCGACCCATGGCAGCGGCACGCTC	90.1
Qy	901	AAGGTTCTTCAAGTGGCCAGAAAGCAACCTGTACAGTTCTACGCCAATCATTAATTTC	96.0
Db	902	AAGGTTCTTCAAGTGGCCAGAAAGCAACCTGTACAGTTCTACGCCAATCATTAATTTC	96.1
Qy	961	ATAACAAGTAGGCCCTGTGATCTCGCGCGCAAACTCACTGTCATCAAGATGCTGTGTAC	102.0
Db	962	ATAACAAGTAGGCCCTGTGATCTCGCGCGCAAACTCACTGTCATCAAGATGCTGTGTAC	102.1
Qy	1021	CAGTATGACAAAATGGAAAACCTATTAAGATATCCCGGTGA	106.2
Db	1022	CAGTATGACAAAATGGAAAACCTATTAAGATATCCCGGTGA	106.3

```

RESULT 8
US-10-174-364-21
Sequence 21, Application US/10174364
Publication No. US20030216308A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-72961P2
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1060)
US-10-174-364-21

Query Match          99.7%; Score 1058.8; DB 15; Length 1343;
Best Local Similarity 99.8%; Pred. No. 2.1e-309;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0

QY      1  ATGAGCAGCAGCAGCCGACCTCCGACGCAACAGCTGCTGTGTGTGTCGCCGCGC 60
Db      2  ATGAGCAGCAGCAGCCGACCCACCTCCGAGCCAAAGCTGCTGTGTGTGTCGCCGCGC 61
QY      61  TCGGCGCTGGGCTTGAGTTTCGTGCCCGGTGCTCTACTAAGCCTCTTGCTGCTCGCT 120

```

Db	62	TGGGCTGCGGCTTGGGTTTCGTGCGCGGTGGTCTACCTACAGCTCTTGCTGTGGCTCGGT	121
OY	121	TTACCAAGAAATATCTTGAACAGTGTATCTCTTCCAGCTGTGTGGCAAGAACAGAG	180
Db	122	TTACCAAGAAATATCTTGAACAGTGTATCTCTTCCAGCTGTGTGGCAAGAACAGAG	181
OY	181	TCCTCTTCACTATCTTGTGGCACTCGCTGTCGCACTCTTGGTCTCTTTTCATA	240
Db	182	TCCTCTTCACTATCTTGTGGCACTCGCTGTCGCACTCTTGGTCTCTTTTCATA	241
OY	241	GTGTTTGTGACTTCTCTGTGTGAAGATTTCATCTTGAACATGACAGATGCTCAGGTCCC	300
Db	242	GTGTTTGTGACTTCTCTGTGTGAAGATTTCATCTTGAACATGACAGATGCTCAGGTCCC	301
OY	301	GACAAAGATCATAAGAGTGTGGAAATTCATATCCATCCACACTCCATATGGATTACTGTA	360
Db	302	GACAAAGATCATAAGAGTGTGGAAATTCATATCCATCCACACTCCATATGGATTACTGTA	361
OY	361	CCGTTAAACCATYGCAGGTATATCGCTGTCTGCGCACCCGCTCAAGTACACACGGTCTCA	420
Db	362	CCGTTAAACCATYGCAGGTATATCGCTGTCTGCGCACCCGCTCAAGTACACACGGTCTCA	421
OY	421	TACCAAGCGCGCACCCGGAAGTCACTTGAAGTGTTCATATCACTGCTTCTGTACACAG	480
Db	422	TACCAAGCGCGCACCCGGAAGTCACTTGAAGTGTTCATATCACTGCTTCTGTACACAG	481
OY	481	ATCCCTTATTAAGTGTGGCCCAACATCTGACCTGAAGACTACATCAGCACTCTGTGAT	540
Db	482	ATCCCTTATTAAGTGTGGCCCAACATCTGACCTGAAGACTACATCAGCACTCTGTGAT	541
OY	541	CAGTCTCATCTGTGATCACTGCTTCAACGCTCTAAGCTGTGTCCCTGCTCATCTTCTTC	600
Db	542	CAGTCTCATCTGTGATCACTGCTTCAACGCTCTAAGCTGTGTCCCTGCTCATCTTCTTC	601
OY	601	ATCTTGAACCTAAATCATTTGTGTGACAAGTCAAGAGGAGAGCAATTTGTCGTCCGTGGC	660
Db	602	ATCTTGAACCTAAATCATTTGTGTGACAAGTCAAGAGGAGCAATTTGTCGTCCGTGGC	661
OY	661	TACTTCAAGGGGAAGACCAACCGCATCTTGTTCACATTAACCTCCATCTTTGSCACACTT	720
Db	662	TACTTCAAGGGGAAGACCAACCGCATCTTGTTCACATTAACCTCCATCTTTGSCACACTT	721
OY	721	TGGGCCCCCGCATCATATGATTTCTTTACACCTCTATGTGGGGCGCCATCTCAAGAACGCG	780
Db	722	TGGGCCCCCGCATCATATGATTTCTTTACACCTCTATGTGGGGCGCCATCTCAAGAACGCG	781
OY	781	TGGCTGTGGACATCATATGTCGACATTTGSCAAATGCTAGCCCTTCTGAACACAGCCATC	840
Db	782	TGGCTGTGGACATCATATGTCGACATTTGSCAAATGCTAGCCCTTCTGAACACAGCCATC	841
OY	841	AACTTCTTCTCTATCTGCTTCATCAGCAAGCGGTTCCGACCATGTGGACGCCACGCTC	900
Db	842	AACTTCTTCTCTATCTGCTTCATCAGCAAGCGGTTCCGACCATGTGGACGCCACGCTC	901
OY	901	AAGGCTTCTTCAAGTGCAGAACCAACCTGTACAGTTCTACACCAATCATTAACCTTTCC	960
Db	902	AAGGCTTCTTCAAGTGCAGAACCAACCTGTACAGTTCTACACCAATCATTAACCTTTCC	961
OY	961	ATTAACAAGTAGCCCCGTGATCTCGCCGGAACCTCACACTGATCAAGATGTGTGTATC	1020
Db	962	ATTAACAAGTAGCCCCGTGATCTCGCCGGAACCTCACACTGATCAAGATGTGTGTATC	1021
OY	1021	CAGTATGACAAAAATGAAAACTATTAAGATATCCCGTGA	1062
Db	1022	CAGTATGACAAAAATGAAAACTATTAAGATATCCCGTGA	1063

```

RESULT 9
US-10-246-583-21
; Sequence 21, Application US/10246583
; Publication No. US2004005862A1
; GENERAL INFORMATION:

```

```

/ APPLICANT: Majmuder
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2CON1
/ CURRENT APPLICATION NUMBER: US/10/246,583
/ PRIOR FILING DATE: 2002-12-06
/ PRIOR APPLICATION NUMBER: 10/174,364
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 128
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 21
/ LENGTH: 1343
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (2)..(1060)
/ US-10-246-583-21

```

```

Query Match      99.7%; Score 1058.8; DB 16; Length 1343;
Best Local Similarity 99.8%; Pred. No. 2.1e-309;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGACACACGACGCGCCACCTCGACAGCAAGCTCGCTCTCTTGTGTCCTCCCGGC 60
DB 2 ATGAGACACACGACGCGCCACCTCGACAGCAAGCTCGCTCTCTTGTGTCCTCCCGGC 61
QY 61 TCGGCTGCGGCTTGGGTTTCGTGCGCGGTCTACTACAGCCTCTTGTGTCCTCGT 120
DB 62 TCGGCTGCGGCTTGGGTTTCGTGCGCGGTCTACTACAGCCTCTTGTGTCCTCGT 121
QY 121 TTACGACAAATATCTTGAAGATGATCCTCTCCAGCTGTGGCAAGAGACAGAG 180
DB 122 TTACGACAAATATCTTGAAGATGATCCTCTCCAGCTGTGGCAAGAGACAGAG 181
QY 181 TCTCTCAACAATATCTTGGACATCGCTGCGGACATCTTGTGTCCTCTTTTCA 240
DB 182 TCTCTCAACAATATCTTGGACATCGCTGCGGACATCTTGTGTCCTCTTTTCA 241
QY 241 GTGTTTGTGACTTCTGTTGGAAGATTTCACTTGAACATGACAGATGCTCAGTCCC 300
DB 242 GTGTTTGTGACTTCTGTTGGAAGATTTCACTTGAACATGACAGATGCTCAGTCCC 301
QY 301 GACAAATCATAGAGTCTGGAATTCATCCATCCACACCTCCATATGATTAAGTGA 360
DB 302 GACAAATCATAGAGTCTGGAATTCATCCATCCACACCTCCATATGATTAAGTGA 361
QY 361 CCGTTAAACATGACAGATATATGCTGTGCGACCGGCTCAAGTACACACGCTCTCA 420
DB 362 CCGTTAAACATGACAGATATATGCTGTGCGACCGGCTCAAGTACACACGCTCTCA 421
QY 421 TACCCAGCCGACCGGAAAGTCAATGTAAGTTTAAATCACTGCTTCTGACAGC 480
DB 422 TACCCAGCCGACCGGAAAGTCAATGTAAGTTTAAATCACTGCTTCTGACAGC 481

```

```

QY 481 ATCCCTATTACTGTTGGCCCAACATCTGGAAGTGAAGTACATGACACCTCTGTGCAT 540
DB 482 ATCCCTATTACTGTTGGCCCAACATCTGGAAGTGAAGTACATGACACCTCTGTGCAT 541
QY 541 CAGTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 542 CAGTCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
QY 601 ACTTGAATCATATCTTGTGTAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 602 ACTTGAATCATATGTTGTGTAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661
QY 661 TACTCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 662 TACTCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
QY 721 TGGGCCCCCGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 722 TGGGCCCCCGCATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
QY 781 TGGCTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 782 TGGCTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
QY 841 AACTTCTCTCTACTGCTTCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 842 AACTTCTCTCTACTGCTTCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
QY 901 AAGCTTTCTTCAAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 902 AAGCTTTCTTCAAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 961
QY 961 ATAAAGATGAGCCCTGATCTGCGCGGCAAACTGACATGATCAAGATGCTGTGATC 1020
DB 962 ATAAAGATGAGCCCTGATCTGCGCGGCAAACTGACATGATCAAGATGCTGTGATC 1021
QY 1021 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGTGA 1062
DB 1022 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGTGA 1063

```

```

RESULT 10
US-10-689-832-21
/ Sequence 21, Application US/10689832
/ Publication No. US20040121380A1
/ GENERAL INFORMATION:
/ APPLICANT: Majmuder, Kamud
/ TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
/ FILE REFERENCE: 15966-729DIV1
/ CURRENT APPLICATION NUMBER: US/10/689,832
/ PRIOR FILING DATE: 2003-10-20
/ PRIOR APPLICATION NUMBER: 09/813,432
/ PRIOR FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78

```

SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
US-10-689-833-21

Query Match 99.7%; Score 1058.8; DB 17; Length 1343;
Best Local Similarity 99.8%; Pred. No. 2.1e-309;
Matches 1060; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
1 ATGAGACACACGACGCGCCACCTCGACGACCAAGCTCGCTGCTTGGTGTGCTCCCGCGC 60
2 ATGAGACACACGACGCGCCACCTCGACGACCAAGCTCGCTGCTTGGTGTGCTCCCGCGC 61
61 TCGGCTGCGGCTTGGGCTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
62 TCGGCTGCGGCTTGGGCTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
121 TTAACGCAAAATCTTGAACAGTATCTCTCCAGTGGTGGCAAGAGAGAGAG 180
122 TTAACGCAAAATCTTGAACAGTATCTCTCCAGTGGTGGCAAGAGAGAGAG 181
181 TCTCTCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
182 TCTCTCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
241 GTGTTGTGAGCTTCTGTTGGAAGATTCATCTTGAACAGTGGTGGCAAGAGAGAG 300
242 GTGTTGTGAGCTTCTGTTGGAAGATTCATCTTGAACAGTGGTGGCAAGAGAGAG 301
301 GACAAATCATGAATGCTGGAATCTCTCATCTCAACCTCCATATGATATCTGTA 360
302 GACAAATCATGAATGCTGGAATCTCTCATCTCAACCTCCATATGATATCTGTA 361
361 CGGTAAACATGACAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
362 CGGTAAACATGACAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
421 TACCCAGCCGCAACCGGAAAGTCAATGTAAGTGTAACTCACTGCTTCTGACAG 480
422 TACCCAGCCGCAACCGGAAAGTCAATGTAAGTGTAACTCACTGCTTCTGACAG 481
481 ATCCCTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
482 ATCCCTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
541 CAGCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
542 CAGCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
601 ATCTTGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
602 ATCTTGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
661 TACTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
662 TACTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
721 TGGGCCCCCGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
722 TGGGCCCCCGGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
781 TGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
782 TGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
841 AACTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
842 AACTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
901 AAGCTTTCTTCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
```

```
Db 902 AAGCTTTCTTCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 961
Qy 961 ATACCAAGTACGCGCTGATCTGCGCGGCAAACTGACATGATGATGATGATGATGATG 1020
Db 962 ATACCAAGTACGCGCTGATCTGCGCGGCAAACTGACATGATGATGATGATGATGATG 1021
Qy 1021 CAGTATGACAAATGAGAAACCTATTAAGATGATGATGATGATGATGATGATGATGAT 1062
Db 1022 CAGTATGACAAATGAGAAACCTATTAAGATGATGATGATGATGATGATGATGATGAT 1063
```

RESULT 11
US-10-450-590-1

Sequence 1, Application US/10450590
Publication No. US20040076985A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
FILE REFERENCE: LIO316 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/450,590
CURRENT FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/254,923
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: US 60/280,110
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: US 60/299,474
PRIOR FILING DATE: 2001-06-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 1059
TYPE: DNA
ORGANISM: Homo sapiens
US-10-450-590-1

Query Match 99.6%; Score 1057.4; DB 16; Length 1059;
Best Local Similarity 99.9%; Pred. No. 4.8e-309;
Matches 1058; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
1 ATGAGACACACGACGCGCCACCTCGACGACCAAGCTCGCTGCTTGGTGTGCTCCCGCGC 60
2 ATGAGACACACGACGCGCCACCTCGACGACCAAGCTCGCTGCTTGGTGTGCTCCCGCGC 61
61 TCGGCTGCGGCTTGGGCTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
62 TCGGCTGCGGCTTGGGCTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
121 TTAACGCAAAATCTTGAACAGTATCTCTCCAGTGGTGGCAAGAGAGAGAGAGAGAGAG 180
122 TTAACGCAAAATCTTGAACAGTATCTCTCCAGTGGTGGCAAGAGAGAGAGAGAGAGAG 181
181 TCTCTCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
182 TCTCTCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
241 GTGTTGTGAGCTTCTGTTGGAAGATTCATCTTGAACAGTGGTGGCAAGAGAGAGAG 300
242 GTGTTGTGAGCTTCTGTTGGAAGATTCATCTTGAACAGTGGTGGCAAGAGAGAGAG 301
301 GACAAATCATGAATGCTGGAATCTCTCATCTCAACCTCCATATGATATCTGTA 360
302 GACAAATCATGAATGCTGGAATCTCTCATCTCAACCTCCATATGATATCTGTA 361
361 CGGTAAACATGACAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
362 CGGTAAACATGACAGGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
421 TACCCAGCCGCAACCGGAAAGTCAATGTAAGTGTAACTCACTGCTTCTGACAG 480
422 TACCCAGCCGCAACCGGAAAGTCAATGTAAGTGTAACTCACTGCTTCTGACAG 481
481 ATCCCTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
```

Dp	1	ATGAGACACACGACGCCCACTTCGACGCAACAGCTGCTGTCTTGGTGTGCCCCGGAC	60
Qy	61	TGGGCTCGGGCTTGGGTTTCGGCCCGNGCTACTACAGCTCTTGTGCTCGCTCGGT	12
Dp	61	TCGGCTCGGGCTTGGGTTTGGTGCCTGGGTCTACAGCTCTTTCAGTGGCTTGGT	12
Qy	121	TTACACGCAAAATCTTGAAGTATCATCTCTTCCAGCTGGGCAAGAACAGAG	180
Dp	121	TTACACGCAAAATCTTGAAGTATCATCTCTTCCAGCTGGGCAAGAACAGAG	180
Qy	181	TCTCTCTAACATCTCTTGGCACTGCTGCTGCGGACATCTTGCTCTTTTTTATA	240
Dp	181	TCTCTCTAACATCTCTTGGCACTGCTGCTGCGGACATCTTGCTCTTTTTTATA	240
Qy	241	GTGTTTGGGACTCTCTGTGGAAATTCATCTTGAACATGACATGAGCTCAGGTCCC	300
Dp	241	GTGTTTGGGACTCTCTGTGGAAATTCATCTTGAACATGAGATGCTCAGGTCCC	300
Qy	301	GACAAATCATAGAAAGTCTGGAAATTTCTCATCCATCCACACTTCATATGATTAATCTGA	360
Dp	301	GACAAATCATAGAAAGTCTGGAAATTTCTCATCCATCCACACTTCATATGATTAATCTGA	360
Qy	361	CGTTAACCATTTGACAGGTATATGCTGTGCGCAACCGGTCAAGTACCAAGGTTCTA	420
Dp	361	CGTTAACCATTTGACAGGTATATGCTGTGCGCAACCGGTCAAGTACCAAGGTTCTA	420
Qy	421	TACCAACCCGCAACCCGGAAAGTCATTTGATAGTGTTAACATCACTGTCTTCTGACACG	480
Dp	421	TACCAACCCGCAACCCGGAAAGTCATTTGATAGTGTTAACATCACTGTCTTCTGACACG	480
Qy	481	ATCCCTATTACTGTGGCCCAACTCTGGAATGGAAGATCAATCAATGCAACCTCTGTGCAT	540
Dp	481	ATCCCTATTACTGTGGCCCAACTCTGGAATGGAAGATCAATCAATGCAACCTCTGTGCAT	540
Qy	541	CAGTCTCATCTGGAATCCATGCTTCAACGTTCACTGGTGCCCTGCTCATCTTCTTC	600
Dp	541	CAGTCTCATCTGGAATCCATGCTTCAACGTTCACTGGTGCCCTGCTCATCTTCTTC	600

Db	661	TACTCCAGGGGAGAGCCACCGCAGCTCTGTTCCATTCATTAACATTAACCTTCATCTTTCGCATTA	720
Qy	721	TGGGCCCCCGGCGATCATCTAATGATTTGTTTACCACTTCTATAGGAGCGCCCATCCAGAACCGC	780
Db	721	TGGGCCCCCGGCGATCATCTAATGATTTGTTTACCACTTCTATAGGAGCGCCCATCCAGAACCGC	780
Qy	781	TGGCTGTGTGCATCATATGTCGACAATTGGCAACATGTCAGCCCTTCTGAAACAGCCATC	840
Db	781	TGGCTGTGTGCATCATATGTCGACAATTGGCAACATGTCAGCCCTTCTGAAACAGCCATC	840
Qy	841	AACTTCTTCTCTACTGTGTCATCGAAGGCGTTCGGCACCATATGGACCGCGCCACGCTC	900
Db	841	AACTTCTTCTCTACTGTGTCATCGAAGGCGTTCGGCACCATATGGACCGCGCCACGCTC	900
Qy	901	AAGGTTTCTTCAAGTGTGCAGAGAACCTGTACAGTTCTACACCAATCATTAACCTTTTTC	960
Db	901	AAGGTTTCTTCAAGTGTGCAGAGAACCTGTACAGTTCTACACCAATCATTAACCTTTTTC	960
Qy	961	ATACAAATTAACCCCTGGATCTGGCGGCAAACTCACATGTGCATCAAGATGCTGTATAC	1020
Db	961	ATACAAATTAACCCCTGGATCTGGCGGCAAACTCACATGTGTCTGTATAC	1020

RESULT 13
US-10-012-140-10

Sequence 10, Application US/10012140
Publication No. US20030009017A1
GENERAL INFORMATION:
APPLICANT: Leiby, Kevin R.
APPLICANT: Kapeller-Lieberman, Rosana
APPLICANT: Gluckerman, Maria A.
TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 381552004900
CURRENT APPLICATION NUMBER: US/10/012,140
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: 60/246,768
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,772
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,185
PRIOR FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 1526
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (139)...(1200)
US-10-012-140-10

Query Match 98.9%; Score 1050.8; DB 14; Length 1526;
Best Local Similarity 99.3%; Pred. No. 5.8e-307; Indels 0; Gaps 0;
Matches 1055; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGAGCAGACGCGACCGCCACCTCGACAGCCAGAGCTCGCTGTTGGTGGTCCCCGGC 60
DB 139 ATGGAGCAGACGCGACCGCCACCTCGACAGCCAGAGCTCGCTGTTGGTGGTCCCCGGC 198
QY 61 TCGGCTGGGCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 120
DB 199 TCGGCTGGGCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 258
QY 121 TTACGAGCAAAATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 180
DB 259 TTACGAGCAAAATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAAGAGAAG 318
QY 181 TCCTCTACAACTATCTTGGCACTCGCTGCTGCGAGACCTTGGTCTCTTTTCAT 240
DB 319 TCCTCTACAACTATCTTGGCACTCGCTGCTGCGAGACCTTGGTCTCTTTTCAT 378
QY 241 GTGTTGTGAGCTTCTGTTGGAAGATTTGATCTTGAACATGACATGCTCAGTCCCGC 300
DB 379 GTGTTGTGAGCTTCTGTTGGAAGATTTGATCTTGAACATGACATGCTCAGTCCCGC 438
QY 301 GACAGATCATAGAGTGTGAATTTCTCATTCATCCACAGCTCCCATATGATTAAGTGA 360
DB 439 GACAGATCATAGAGTGTGAATTTCTCATTCATCCACAGCTCCCATATGATTAAGTGA 498
QY 361 CCGTTAACAATGACAGATATCGCTGTCTGACACCGGCTCAAGTACCAACAGGCTCTCA 420
DB 499 CCGTTAACAATGACAGATATCGCTGTCTGACACCGGCTCAAGTACCAACAGGCTCTCA 558
QY 421 TACCCGCGCGACCGCGAAAGTCAATTTGATGTTTATCATCACTGCTCTGACACAGC 480
DB 559 TACCCGCGCGACCGCGAAAGTCAATTTGATGTTTATCATCACTGCTCTGACACAGC 618
QY 481 ATCCCTATTAATGAGTGGCGCAACATCTGGAAGTGAAGATCAATCAAGACCTGATGAT 540
DB 619 ATCCCTATTAATGAGTGGCGCAACATCTGGAAGTGAAGATCAATCAAGACCTGATGAT 678
QY 541 CAGGTCTCATCTGATTCACCTGCTTCAACCGTCTACCTGCTGCTGCTGCTGCTGCTGCT 600
DB 679 CAGGTCTCATCTGATTCACCTGCTTCAACCGTCTACCTGCTGCTGCTGCTGCTGCTGCT 728

QY 601 ATCTGACCTCATCTATGTTGTAACAGCTCAGGAGGAAGCAATTTGGTCTCCGATGC 660
DB 739 ATCTGACCTCATCTATGTTGTAACAGCTCAGGAGGAAGCAATTTGGTCTCCGATGC 798
QY 661 TACTCCAGGGGGAAGACCAACCGCATCTTGTTCACCATTAACCTCACTTTGGCACACTT 720
DB 799 TACTCCAGGGGGAAGACCAACCGCATCTTGTTCACCATTAACCTCACTTTGGCACACTT 858
QY 721 TGGGCCCCCGCATCATATGATTTTAAACACCTTAAAGGGGCGCCCATCCGAACCGC 780
DB 859 TGGGCCCCCGCATCATATGATTTTAAACACCTTAAAGGGGCGCCCATCCGAACCGC 918
QY 781 TGGCTGTGACATCATATGTTCCGACATTTGCAACATGCTAGCCCTTGAACAGACCATC 840
DB 919 TGGCTGTGACATCATATGTTCCGACATTTGCAACATGCTAGCCCTTGAACAGACCATC 978
QY 841 AACTCTTCTCTACTGCTTCAATCAAGCAAGCGGTTCCGACATGACAGCGCGACGCTC 900
DB 979 AACTCTTCTCTACTGCTTCAATCAAGCAAGCGGTTCCGACATGACAGCGCGACGCTC 1038
QY 901 AAGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTTACACCATCAATCACTTTCC 960
DB 1039 AAGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTTACACCATCAATCACTTTCC 1098
QY 961 ATTAACAAGTACCGCCGATCTCGCGGCAAACTCACTGATCAAGATGCTGTGTAC 1020
DB 1099 ATTAACAAGTACCGCCGATCTCGCGGCAAACTCACTGATCAAGATGCTGTGTAC 1158
QY 1021 CAGTATGACAAAATGGAAGAACTTAAAGTATCCCGTGA 1062
DB 1159 CAGTATGACAAAATGGAAGAACTTAAAGTATCCCGTGA 1200

RESULT 14
US-09-432-19
Sequence 19, Application US/09813432
Publication No. US20030148485A1
GENERAL INFORMATION:
APPLICANT: Taupier Jr., Raymond J
APPLICANT: Majumder, Kamud
APPLICANT: Spaderna, Steven K
APPLICANT: Smithson, Glenda
APPLICANT: Mezes, Peter S
TITLE OF INVENTION: Verne, Corine A. M.
FILE REFERENCE: 15966-729
CURRENT APPLICATION NUMBER: US/09/813,432
CURRENT FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/193,843
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19

OY	961	ATAACAGTGGCCCTGATCTGCCGGCGAACTCAACTGTGCATCAAGATGCTGTGTAC	1020
Dd	961	ATAACAGTAGTCCCCTGACTCTGCCGGCGAACTCAACTGTGCATCAAGATGCTGTGTAC	1020
OY	1021	CAGTGATGCAAAAATGGAAAAAACCATTAAAA	1050
Dd	1021	CAGTGATGCAAAAATGGAAAAAACCATTAAAA	1050

```

RESULT 15
US-10-174-364-19
; Sequence 19, Application US/10174364
; Publication No. US20030216308A1
GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2
; CURRENT APPLICATION NUMBER: US/10/174,364
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1116)
US-10-174-364-19
Query Match 98.7%; Score 1048.4; DB 15; Length 1119;
Best Local Similarity 99.9%; Prod. No. 2.ee-306;
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGAGCAGCAGCAGCGCCACCTTCGAGCCCAAGCTCGCTGTCCTGGTGTGCTCCCGGCG 60
1 ATGAGCAGCAGCAGCGCCACCTTCGAGCCCAAGCTCGCTGTCCTGGTGTGCTCCCGGCG 60
1 TCGGCCCTGGCGGCTTGGGTTTCGTCGCCCGGTGCTTACTAAGAGCTCTTGTCGTCCGAT 120
61 TCGGCCCTGGCGGCTTGGGTTTCGTCGCCCGGTGCTTACTAAGAGCTCTTGTCGTCCGAT 120
121 TTACAGCAAAATATCTTGGACAGTATCATCTCTCCACGTGGTGGCAAGAGACAAG 180
121 TTACAGCAAAATATCTTGGACAGTATCATCTCTCCACGTGGTGGCAAGAGACAAG 180
181 TCCTCTTAACAATATCTTGGACAGTATCATCTCTCCACGTGGTGGCAAGAGACAAG 240
181 TCCTCTTAACAATATCTTGGACAGTATCATCTCTCCACGTGGTGGCAAGAGACAAG 240
241 GTGTTGTGAGCTTCTGTTGGAGATTTCATCTTGAACATTCAGATATGCTTATGCTTCTT 240

```


Db 241 GTGTTGTGACTTCCTGTTGGAAGATTTCATCTTGAACATGACAGATGCCTCAGTCC 300
QY 301 GACAGATCATAGAGTGTGGAATTCATCATCCATCCACACTCCATATGATTAAGTGA 360
Db 301 GACAGATCATAGAGTGTGGAATTCATCATCCATCCACACTCCATATGATTAAGTGA 360
QY 361 CGGTAAACCATGACAGGTATATCGGTCTGACACCGGCTCAAGTAAACAAGGTCTCA 420
Db 361 CGGTAAACCATGACAGGTATATCGGTCTGACACCGGCTCAAGTAAACAAGGTCTCA 420
QY 421 TACCCAGCCCGCACCCGGAAGTCATTTGAAGTTTACATCACTGCTTCTGACCAGC 480
Db 421 TACCCAGCCCGCACCCGGAAGTCATTTGAAGTTTACATCACTGCTTCTGACCAGC 480
QY 481 ATCCCTATTACTGTGGGCCCAATCTGGACTGAAGACTACATCAGACACTCTGTGCAT 540
Db 481 ATCCCTATTACTGTGGGCCCAATCTGGACTGAAGACTACATCAGACACTCTGTGCAT 540
QY 541 CAGTCTCATCTGGATCCACCTGCTTCAACCTGCTACCTGGCCCTGCTCACTTCTTC 600
Db 541 CAGTCTCATCTGGATCCACCTGCTTCAACCTGCTACCTGGCCCTGCTCACTTCTTC 600
QY 601 ATCTGAACATCATCTGTTGTACAAGCTCAGAGGAAAGCAATTTTCTCCGTGGC 660
Db 601 ATCTGAACATCATCTGTTGTACAAGCTCAGAGGAAAGCAATTTTCTCCGTGGC 660
QY 661 TACTCCAGGGGAAAGCAACCGGCATCTTGTTCACCATTAACCTTCTTGGCACATT 720
Db 661 TACTCCAGGGGAAAGCAACCGGCATCTTGTTCACCATTAACCTTCTTGGCACATT 720
QY 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTCATATGGGGGCGCCATCCAGAACCG 780
Db 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTCATATGGGGGCGCCATCCAGAACCG 780
QY 781 TGGCTGTGACATCATGTTCGACATTTGCCAATGCTAGGCCCTTCTGAACAGCCATC 840
Db 781 TGGCTGTGACATCATGTTCGACATTTGCCAATGCTAGGCCCTTCTGAACAGCCATC 840
QY 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGGCAGCGCCACGCTC 900
Db 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGGCAGCGCCACGCTC 900
QY 901 AAGGCTTTCTTCAAGGCCAGAGCAACCTGTACAGTTCTPACCAATCATTAATTTC 960
Db 901 AAGGCTTTCTTCAAGGCCAGAGCAACCTGTACAGTTCTPACCAATCATTAATTTC 960
QY 961 ATTAACAAGTAGCCCTGATCTGCGCGCAACTCACACTGCATCAAGATGCTGTGTAC 1020
Db 961 ATTAACAAGTAGCCCTGATCTGCGCGCAACTCACACTGCATCAAGATGCTGTGTAC 1020
QY 1021 CAGTATGACAAAAATGAAAACTTATAAA 1050
Db 1021 CAGTATGACAAAAATGAAAACTTATAAA 1050

Search completed: December 15, 2004, 02:53:10
Job time : 544.543 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:17:59 ; Search time 3344.6 Seconds
(without alignments)
11570.599 Million cell updates/sec

Title: US-09-995-225B-15

Perfect score: 1062
Sequence: 1 atggagacacagcagcagccca.....ctataaagatcccgaga 1062

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	469.8	44.2	803	9	CC905917 t025011ba
C 2	262	24.7	362	8	AZ513631 IM0359L07
C 3	241.4	22.7	448	8	AQ225693 HS 2009. B
C 4	218.8	20.6	813	8	B2169352 CH230-259
C 5	209.8	19.8	806	9	CNS02284 Tetradon
C 6	200	18.8	588	8	AQ344035 RBC11-12
C 7	187	17.7	603	8	AQ314433 RBC11-11
C 8	187.6	17.7	947	9	AL190367 Tetradon
C 9	185	17.4	708	5	BU610327 UI-M-DJ2-
C 10	183	17.2	824	9	CL137737 ISB1-110A
C 11	172.4	16.2	1269	9	CL649540 CH213-237
C 12	150	14.1	867	9	CL137848 ISB1-110C
C 13	149	14.0	898	9	AL193966 Tetradon
C 14	112.6	10.6	515	1	AL921815 AL921815
C 15	111.8	10.5	1107	9	CNS0417F Tetradon
C 16	87.6	8.2	816	8	AZ535744 ENTQ25TR
C 17	81.6	7.7	891	8	AZ683582 ENTQ47TR
C 18	81.6	7.7	903	8	BH153606 ENTQ83TR
C 19	77.8	7.3	843	8	AZ551618 ENTQV54TR
C 20	77.4	7.3	908	8	AZ548467 ENTQK30TR
C 21	75	7.1	900	8	BH146886 ENTQK48TR
C 22	71.8	6.8	900	8	AZ549980 ENTQD94TR
C 23	71	6.7	976	8	BH149983 ENTQD93TR
C 24	70.2	6.6	358	2	BB870915 BB870915

C 25	69.2	6.5	877	8	AZ531291
C 26	69.2	6.5	912	8	AZ551092
C 27	67.8	6.4	488	9	FR0012324
C 28	67.6	6.4	888	8	AZ528430
C 29	67.4	6.3	849	8	AZ546009
C 30	67	6.3	905	8	AZ550256
C 31	66.2	6.2	721	5	BW141179
C 32	66	6.2	931	8	BH160272
C 33	63.2	6.0	890	8	AZ530768
C 34	61.8	5.8	774	9	CNS02APC
C 35	61.8	5.8	881	7	CNS97894
C 36	61.4	5.8	607	4	BJ348812
C 37	61.2	5.8	692	7	CR437869
C 38	61.2	5.8	726	7	CR438428
C 39	61.2	5.8	735	7	CR441503
C 40	61.2	5.8	768	7	CR445667
C 41	60.8	5.7	880	8	AZ529191
C 42	60.8	5.7	954	7	CO250025
C 43	60.6	5.7	712	7	CR438462
C 44	60	5.6	787	5	BW391073
C 45	60	5.6	1023	9	CNS05805

ALIGNMENTS

RESULT 1
CC905917/c 803 bp DNA linear GSS 08-AUG-2003
LOCUS t025011ba.r1 TAMBT Bos taurus genomic clone t025011ba, genomic
DEFINITION survey sequence.
ACCESSION CC905917
VERSION CC905917.1 GI:33524850
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 803)
Lin.S., Najjar,F.Z., Adelson,D., Gilli,C.A. and Roe,B.A.
Bovine BAC End Sequences from Library TAMBT
JOURNAL Unpublished (2003)
COMMENT Contact: Bruce A. Roe
Advanced Center for Genome Technology
University of Oklahoma Department of Chemistry and Biochemistry
620 Parrington Oval, Room 208, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Classes: BAC ends
High quality sequence start: 39
High quality sequence stop: 551.
Location/Qualifiers
1..803
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Angus bull T A M U Shoshone Y6 11519666"
/db_xref="taxon:9913"
/clone="t025011ba"
/sex="Male"
/cell_type="Blood"
/clone_lib="TAMBT"
/note="Vector: pBelOBAcl1, site_1: HindIII, site_2: HindIII; TAMBT Bovine BAC library (Male) produced by Texas A&M University, Department of Animal Science."

FEATURES

ORIGIN

Query Match 44.2%; Score 469.8; DB 9; Length 803;
Best Local Similarity 89.2%; Pred. No. 4.2e-114;
Matches 518; Conservative 0; Mismatches 62; Indels 1; Gaps 1;

QY 483 CCCCTATTACTGTGGCCCAACATCTGAGCTAGACTAGACTACATCAGAC-CTCTGTGATC 541

```

Db      766  CCTTAAAGCTGGGGGCCCCACCCCTGAGCTTGAATTCATCAGACCGGTCCATGATTC
Qy      542  AGCTCTCATCTGGATTCACCTGCTTACCGGTCTACCTGGTCCCTGCTCATCTTCTCA
Db      706  ATGTCTCTTATCTGAATTCACCTGCTTCCGTGTGATCGGTACCTCTTCATATTTCTCA
Qy      602  TCTTGAATCATCATTTGTGTCAAGCTCAGAGAGAGAGCAATTTTGTCTCGTGTGCT
Db      646  TCTGAATCTCATCATTTGTGTCAAGCTCAGAGAGAGAGCAATTTTGTCTCGTGTGCT
Qy      662  ACTTCACGGGGAGAGACCAACCGCATCTTGTGTCAACATTACCTCCATCTTTGCCACACTT
Db      586  ACTTCACGGGGAGAGAGACCAACCGCATCTTGTGTCAACATTACCTCCATCTTTGCCACACTT
Qy      722  GGGCCCCCGCATCATCATTTTCTTTCACACTCTTATGGGCGCCCATTCAGAACCGCT
Db      526  GGGCCCCCGCATCATCATTTTCTTTCACACTCTTATGGGCGCCCATTCAGAACCGCT
Qy      782  GGCTGTGTGACATCATCTGTCCGACATTTGGCAACATGTCGCTTGTGAACACGGCATCA
Db      466  GGCTGTGTGACATCATCTGTCCGACATTTGGCAACATGTCGCTTGTGAACACGGCATCA
Qy      842  ACTTCTCTCTACCTGCTTTCATCAGCAAGCGGTTCCGACCATGGCAGCGCCACGCTCA
Db      406  ACTTCTCTCTACCTGCTTTCATCAGCAAGCGGTTCCGACCATGGCAGCGCCACGCTCA
Qy      902  AGGCTTTCTTCAAGTCCAGAGCAACCGTTCATGACCAATCATCAATCAATTTTCCA
Db      346  AGGCTTTCTTCAAGTCCAGAGCAACCGTTCATGACCAATCATCAATCAATTTTCCA
Qy      962  TAAAGATGAGCCCTGATCTCGCGGCAAACTCACACTGCATCAAGATCTGTGTAC
Db      286  TAAAGATGAGCCCTGATCTCGCGGCAAACTCACACTGCATCAAGATCTGTGTAC
Qy      1022  AGTATGACAAATAATGAAAACTATTAAGTATCCCGTCA
Db      226  AGTATGACAAATAATGAAAACTATTAAGTATCCCGTCA

```

```

RESULT 2
LOCUS      A2513631/c      362 bp      DNA      linear      GSS 05-OCT-2000
DEFINITION 1M0359L07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
ACCESSION  A2513631
VERSION    A2513631.1
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 362)
            Dum,D. Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D., Weis,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0359 row: L column: 07
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 362.

```

FEATURES

SOURCE

```

Location/Qualifiers
1..362
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0359L07"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g1473214[gbl]AR129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

ORIGIN

```

Query Match      24.7%; Score 262; DB 8; Length 362;
Best Local Similarity 91.4%; Pred. No. 1e-58;
Matches 277; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy      760  GGGGGCCCATCCAGAACCCCTGCTGTGCAATCATGTCGACATTTGCCAATGCTTA
Db      361  GGAGACCCCATCCAGAACCCCTGCTGTGCAATCATGTCGACATTTGCCAATGCTTA
Qy      820  GCCCTTGTGAACACAGGCATCACTTCTCTCTATCTGCTTATCGAAGGGTCCGC
Db      301  GCCCTTGTGAACACAGGCATCACTTCTCTCTATCTGCTTATCGAAGGGTCCGC
Qy      880  ACCATGGCAGCGCCACGCTCAAGGCTTCTTCAAGTCCAGAACCACTTGTAGAGTTT
Db      241  ACCATGGCAGCTCCACACTCAAGGCTTCTTCAAGTCCAGAACCACTTGTAGAGTTT
Qy      940  TACACCAATCATTAATCTTCCATTAACAAGTAGCCCTGGATCTCGCCGCAAACTCAC
Db      181  TATACCAACCATTAATCTTCCATTAACAAGTAGCTCTGGATCTCGCCGCAAACTCAC
Qy      1000  TGCATGAAGATGCTGTGTGTACAGATGTGACAAATAATGAAAACTTATTAAGTATCCCG
Db      121  TGCATGAAGATGCTGTGTGTACAGATGTGACAAATAATGAAAACTTATTAAGTATCCCG
Qy      1060  TGA 1062
Db      61  TGA 59

```

```

RESULT 3
LOCUS      A0225693/c      448 bp      DNA      linear      GSS 26-SEP-1998
DEFINITION HS 2009 B2 B08 T7 CIT Approved Human Genomic Sperm Library D Homo
ACCESSION  A0225693
VERSION    A0225693.1
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

REFERENCE 1 (bases 1 to 448)
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Kellier, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D., and Hood, L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2009 row: D column: 16
Class: BAC ends
High quality sequence stop: 448.
Location/Qualifiers
source 1..448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2009 Col=16 Row=D"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"
ORIGIN
Query Match 22.7%; Score 241.4; DB 8; Length 448;
Best Local Similarity 93.7%; Pred. No. 3.5e-53;
Matches 251; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 795 CATGTCGACATTCGACATGCTAGCCCTTGAACACAGCCATCACTTCTCCCTCA 854
DB 445 CAGGTGACATTCGACATGCTAGCCCTTGAACACAGCCATCACTTCTCCCTCA 386
QY 855 CTGCTTCATCAGCAGAGGGTTCGACACATGAGCGCGCCAGCTCAAGGCTTTCTTCA 914
DB 385 CTGCGTCATCAGCAGAGGGTTCGACACATGAGCGCGCCAGCTCAAGGCTTTCTTCA 326
QY 915 GTGCCGAGACACCTGTACAGTTCTACACCAATCAATCTTTCCATAAAGTACCC 974
DB 325 GTGCCGAGACACCTGTACAGTTCTACACCAATCAATCTTTCCATAAAGTACCC 266
QY 975 CTGGAATCTGCGCGGCAACCTCACATGATCAAGATGCTGTACCAAGTATGACAAAA 1034
DB 265 CTGGAATCTGCGCGGCAACCTCACATGATCAAGATGCTGTACCAAGTATGACAAAA 206
QY 1035 TGGAAAACCTATAAAGTATCCCGTGA 1062
DB 205 TGGAAAACCTATAAAGTATCCCGTGA 178
RESULT 4
B2169352 813 bp DNA linear GSS 11-OCT-2002
LOCUS B2169352
DEFINITION CH230-259E22, TV CHORI-230 Segment 2, Rattus norvegicus genomic clone
ACCESSION CH230-259E22, genomic survey sequence.
VERSION B2169352.1 GI:23810547
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
1 (bases 1 to 813)
Zhao, S., Shetty, J., Shatman, S., Teegave, G., Geer, K., Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,

TIGR, F., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 MbOI segment
Unpublished (1999)
COMMENT CH230-259E22.TV
Other_GSSs: CH230-259E22.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@tigr.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or eting_information.html). BAC end
page: http://www.tigr.org/tldb/bac_end/rat/bac_end_intro.html
Plate: 259 row: E column: 22
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
source 1..813
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-259E22"
/sex="female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"
ORIGIN
Query Match 20.6%; Score 218.8; DB 8; Length 813;
Best Local Similarity 86.7%; Pred. No. 4.5e-47;
Matches 241; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 125 CAGCAAAATCTTTGACAGTATCATCTCTCCAGCTGTGGCAGAGACAGAACTCT 184
DB 536 CAGCAAAATCTTTGACAGTATCATCTCTCCAGCTGTGGCAGAGACAGAACTCT 595
QY 185 CCTACAACTATCTCTGGAAGTCTGCGCGGACATCTTGCTCTTTTATAGTGT 244
DB 596 CCTACAACTATCTCTGGAAGTCTGCGCGGACATCTTGCTCTTTTATAGTGT 655
QY 245 TTGTGACTTCTCTGGAAGTCTGCGCGGACATCTTGCTCTTTTATAGTGT 304
DB 656 TTGTGACTTCTCTGGAAGTCTGCGCGGACATCTTGCTCTTTTATAGTGT 715
QY 305 AGATCATAGAGTCTGGAATTCATCATCATCATCATCATCATCATCATCATCATCAT 364
DB 716 AGATCATAGAGTCTGGAATTCATCATCATCATCATCATCATCATCATCATCATCAT 775
QY 365 TAACCATGACAGTATATCGCTGTGACACCGCTC 402
DB 776 TAACCATGACAGTATATCGCTGTGACACCGCTC 813
RESULT 5
CNS02284 806 bp DNA linear GSS 01-SEP-2000
LOCUS CNS02284
DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone
227D19 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION CNS02284
VERSION AL178069.1 GI:7816126
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

REFERENCE
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

AUTHORS
Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fzames, C., Winker, P., Brothier, P., Quetier, F.,
Saurin, W., and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
20296633
10835645

REFERENCE
2
Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A., and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

JOURNAL
MEDLINE
PUBMED
20359837
10899143

REFERENCE
3
(bases 1 to 806)

AUTHORS
Genoscope.

TITLE
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
location/Qualifiers

FEATURES
source
1..806
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="227D19"
/clone_1lb="G"
/note="Genoscope sequence ID : COAG227DB10LPI-end : T7"

ORIGIN
Query Match 19.8%; Score 209.8; DB 9; Length 806;
Best Local Similarity 56.3%; Pred. No. 1,1e-44;
Matches 418; Conservative 7; Mismatches 289; Indels 29; Gaps 1;

QY 166 GCAAGAAGACAGAGGCTCTCTCACTATCTTTGGCACTGCTGCGCCGACATCTTG 225
Db 2 GGAAGACAGAGAGGCGCTGTACTACTGCGCGCGGTGACAGGCTCCGACATCTTC 61
QY 226 GTCTCTTTTCAATAGTGTGTTGACTTCTGTTGGAAGATTTCATCTGGAACATGAG 285
Db 62 TCCACACTCTTCACTATCTTCGAGGCTTCCTGTTGAGAGCGCGGTTTCCACCGGAG 121
QY 286 ATGCGCTCAGGTCCCGACAGATCATAGAGTCTGGAATTCTCATCTCCACACTCC 345
Db 122 GTCCCGCGCTCTTACACTGACAGCGCGCGAGTTGCGCCACACGAGCCCTCC 181
QY 346 ATATGATTACTGTAACCGTTAACCATTGACAGTATATGCTGCTGTCACCGCTCAAG 405
Db 182 ATCTGTGTCACCGTCCCTTCAACCGTGAACGCTGAGCGCTGTGCAACCCCTCTC 241
QY 406 TACCAACAGGTCTATATCCAGCCCGACCCGGAAGTCATGTGAAGTTTATCATCACC 465
Db 242 CACAGGACAGATCAAGTACCGCGCGGACAGAGAGATCATCGGAGGTCTGCTGCTG 301
QY 466 TCTTCTCTGACAGCATCCCTATTAATGAGGCGCAACATCTGAGTGAAGACTATC 525
Db 302 TCGGCGGCTCGGCGCTCTCTTCTGAGGTCCAGCATGTGAGAAACAGCACCCG 361
QY 526 AGCACTTGTGACATCAAGTCTCATCTGATTCACCTGCTTCAACCGCTACCTGAGCC 585
Db 362 CGAGGCGGTGACGCGCTCTCATCTGAGCCAGAGTGAACATCATCTTCTGCGCC 421
QY 586 TGTCTCATCTTCTCATCTTGAATCAATATTTGTGTAACAGCTCAGAGAGAGCAAT 645

Db 422 TGAGCAATCTTCTGAGTCTTCAACTTTGATCATCAACGCTGAGGTGCGGACARG 481
QY 646 TTTCGTCTCCGTGCTACTTCCAC-----GGGAGACA 676
Db 482 CGGACAGGAGCGGGCGGAGCGCGGCMCAAGTMCACGCTTCGCGCGCTGGGGAGA 541
QY 677 CCACCGCCATCTTTGTTACCATTTACTTCCTGTTGACACTTTTGAGCCCGCCGATCA 736
Db 542 GCAGGCGCATGTGCTGCGCATCCTCGTCTTCTGTGCTGTGGCACCCAGGAGCG 601
QY 737 TCATGATCTTTACACCTCTATGAGGCGCCATCCAGACCGCTGCTGTGACATCA 796
Db 602 TGTGTGATCTACACCTGTACGTTGAGTTTCAAGCGAGMTGGCGCTCCACCTGG 661
QY 797 TGTCCGACATTTGCCAATGCTAGGCTTGTGAACACAGCCATCACTTTCTCTACT 856
Db 662 CCTACGACCTGTCCAAAGCTGCGCATGCTCAACAGCGCGTAACTTTCTCTGACT 721
QY 857 GCTTCATGCAAGGCGGTTCGCG 879
Db 722 GCTTGTACAGCAAGCGCTTCGCG 744

RESULT 6
AQ344035/C 588 bp DNA linear GSS 07-MAY-1999
LOCUS
DECINTITION
PCII11-124J23.TV PCII-11 Homo sapiens genomic clone PCII-11-124J23,
genomic survey sequence.
ACCESSION
AQ344035
VERSION
AQ344035.1 GI:4168931
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 588)
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End Sequences from Library PCII-11 for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@igf.org
Library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.
FEATURES
source
1..588
Location/Qualifiers
1..588
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7547470"
/db_xref="taxon:9606"
/clone="PCII-11-124J23"
/sex="Male"
/cell_type="Lymphocytes"
/clone_1lb="PCII-11"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
PCII11 Human Male BAC Library"

ORIGIN
Query Match 18.8%; Score 200; DB 8; Length 588;
Best Local Similarity 100.0%; Pred. No. 4.3e-42;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 CAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAGAAGTCTT 184
DB 201 CAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAGAAGTCTT 142

QY 185 CCTACAACTATCTCTTGACAGTCTGGTGGTCCGACATCTTGATCTTTTTCATAGTGT 244
DB 141 CCTACAACTATCTCTTGACAGTCTGGTGGTCCGACATCTTGATCTTTTTCATAGTGT 82

QY 245 TTGTGACCTCTCTTGGAAGATTTTCATCTTGAACATGAGATGCTCAGTCCCGACA 304
DB 81 TTGTGACCTCTCTTGGAAGATTTTCATCTTGAACATGAGATGCTCAGTCCCGACA 22

QY 305 AGATCATAGAGTCTGGAA 324
DB 21 AGATCATAGAGTCTGGAA 2

RESULT 7
AQ351433/c 603 bp DNA linear GSS 07-MAY-1999
LOCUS RPI11-113113.TJ RPI1-11 Homo sapiens genomic clone RPI1-113113,
DEFINITION genomic survey sequence.
ACCESSION AQ351433
VERSION AQ351433.1 GI:4178768
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 603)
Zhaio,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.
AUTHORS Use of BAC End Sequences from Library RPI1-11 for Sequence-Ready
Map Building
TITLE Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: RPI11-113113.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeel@igf.org
Clones are derived from the human BAC library RPI1-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (http://resgen.com). BAC end search page:
http://www.igf.org/tldb/humgen/bac_end_search/bac_end_search.html
Seq Primer: SP6
Class: BAC ends.

FEATURES
Source Location/Qualifiers
1..603
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7543212"
/db_xref="taxon:9606"
/clone="RPI1-11-113113"
/sex="Male"
/cell_type="lymphocytes"
/clone_lib="RPI1-11"
/note="Vector: PBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
RPI11 Human Male BAC Library"

ORIGIN
Query Match 18.5%; Score 197; DB 8; Length 603;
Best Local Similarity 100.0%; Pred. No. 2.7e-41;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
125 CAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAGAAGTCTT 184

DB 197 CAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAGAAGTCTT 138
QY 185 CCTACAACTATCTCTTGACAGTCTGGTGGTCCGACATCTTGATCTTTTTCATAGTGT 244
DB 137 CCTACAACTATCTCTTGACAGTCTGGTGGTCCGACATCTTGATCTTTTTCATAGTGT 78

QY 245 TTGTGACCTCTCTTGGAAGATTTTCATCTTGAACATGAGATGCTCAGTCCCGACA 304
DB 77 TTGTGACCTCTCTTGGAAGATTTTCATCTTGAACATGAGATGCTCAGTCCCGACA 18

QY 305 AGATCATAGAGTCTG 321
DB 17 AGATCATAGAGTCTG 1

RESULT 8
CNS02BVO 947 bp DNA linear GSS 01-SEP-2000
LOCUS CNS02BVO
DEFINITION Tetradon nigroviridis genome survey sequence pUC-ori end of clone
254A20 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL190367
VERSION AL190367.1 GI:7828471
KEYWORDS GSS: genome survey sequence.
SOURCE Tetradon nigroviridis
ORGANISM Tetradon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1
Roest Crolius,H., Bailion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fitzmes,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetradon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE
PUBMED 2029633
10835645
2

TITLE Roest Crolius,H., Bailion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Fitzmes,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE
PUBMED 20359837
10899143
3 (bases 1 to 947)

REFERENCE Genoscope.
DIRECT Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.

FEATURES
Source Location/Qualifiers
1..947
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="254A20"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG254BA10SP1-end :
pUC-ori"

ORIGIN
Query Match 17.7%; Score 187.6; DB 9; Length 947;
Best Local Similarity 70.0%; Pred. No. 1e-38;
Matches 266; Conservative 3; Mismatches 102; Indels 9; Gaps 1;
125 CAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAGACAGAAGTCTT 184

```

QY 692 TCACCATCTCTCATCTTTGGCAGACTTTGGGCCCCCGCATCATCATGATCTTAC 751
Db 7 TGGCATACCTCCATCTTTTGGCGTTCTGTGGGCGCCGCGACCTGTATGATCTCTAC 66
QY 752 ACCCTGATGGGGCGCCCATTCAGAACCGCT-----GGCTGTGCATCATGTCG 802
Db 67 ACTTTACTCTGCGCTCCGCGAGCGCTCGGGGGCGCGCGGCTGTGTGACATTTCTAC 126
QY 803 ACATGGCAACATGCTAGCCCTTTCAGACACAGCATATCACTTCTTCTTACTGCTTCA 862
Db 127 ATCTTCCCAACATGTTGGCGTTGCTTCAACACGGGTGTTACTTCTTCTTACTGTTCA 186
QY 863 TCAGCAACGGGTTTCGACCATGTCAGCGCCGACGCTTCAGAGCTTCTTCAAGGCCAGA 922
Db 187 TCAGCAACGGGTTTCGCGGCGCATGGCGGCAACGTGCTCCGAGCGCTCTCCACTGCA 246
QY 923 AGCAACCTGTACAGTTCTTACACCAATCACTTTTTCATTAACAAGTACGCCCTGATCT 982
Db 247 AGCAGTCGACGCGCTTTCACGCGCAGCCACCACTTTTCATCAAGTATGCCCTGATCT 306
QY 983 CGCCGGCAAACTCACACTGATCATGAAGATGCTGTGTACCACTATGACAAAAATGAAAC 1042
Db 307 CACAGCAAACTCCCATGATCATGAAGATGCTGTGTACCACTATGACAAAAATGAAAC 366
QY 1043 CTATTAAGATATCCCGCTGA 1062
Db 367 CCATCTGTATTTCTCTCTGA 386

```

```

RESULT 9          708 bp      mRNA      linear      EST 20-FEB-2003
BU610327          NIH BMAP D12 Mus musculus cDNA clone
LOCUS             BU610327
DEFINITION        BU610327.1 GI:23276542
ACCESSION         BU610327
VERSION           EST
KEYWORDS          Mus musculus (house mouse)
SOURCE            Mus musculus
ORGANISM          Mus musculus
REFERENCE          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLES            Bernaldo, M.F., Lennon, G. and Soares, M.B.
                  Normalization and subtraction: two approaches to facilitate gene
                  discovery
JOURNAL            Genome Res. 6 (9), 791-806 (1996)
MEDLINE            97044477
PUBMED            8889548
COMMENT            Contact: Chin, H
                  National Institute of Mental Health
                  6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
                  20892-9643, USA
                  Tel: 301 443 1706
                  Fax: 301 443 9890
                  Email: mestr@mail.nih.gov
                  Tissue Procurement: Dr. Robin Davisson
                  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
                  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                  Clone Distribution: Researchers may obtain clones from Research
                  Genetics (www.reagen.com).
                  Seg primer: M13 REVERSE.

```

FEATURES

```

source
1..708
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-D12-bw1-f-07-0-UI"
/tissue_type="subfornical organ and postrema"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_id="NIH_BMAP_D12"

```

ORIGIN

```

Query Match          17.4%; Score 185; DB 5; Length 708;
Best Local Similarity 92.8%; Pred.No. 4.6e-38;
Matches 194; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

/note="Organ: Brain; Vector: pRTT3-Pac (Pharmacia) with a modified polylinker. Site 1: EcoR I; Site 2: Not I; UI-M-D12 is a subcloned cDNA library constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRTT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is: GCTACATGAT, subfornical organ and area postrema."

```

QY 854 ACTGCTTATCAGCAAGCGGTCCGACCATGAGCGCCGACGCTCAAGGCTTCTTCA 913
Db 1 ACTGCTTATCAGCAAGCGGTCCGACCATGAGCGCTCAAGGCTTCTTCA 60
QY 914 AGTCCAGAAAGCACTTGAAGTTTACCAATCATTAATCTTTCATTAACAAGTACC 973
Db 61 AGTCCAGAAAGCACTTGAAGTTTACCAATCATTAATCTTTCATTAACAAGTACC 120
QY 974 CCGATCTCGCCGCAAACTCACAAGTATGATGCTGTGTACCAAGTATGACAAA 1033
Db 121 CCGATCTCGCCGCAAACTCACAAGTATGATGCTGTGTACCAAGTATGACAAA 180
QY 1034 ATGAAACCTATTAAGTATCCCGTGA 1062
Db 181 ATGAAACCTATTAAGTATCCCGTGA 209

```

```

RESULT 10
CL137737
LOCUS             CL137737
DEFINITION        CL137737.1 ISB1 Xenopus tropicalis genomic clone ISB1-110A2,
                  genomic survey sequence.
ACCESSION         CL137737
VERSION           GSS
KEYWORDS          Xenopus tropicalis (western clawed frog)
SOURCE            Xenopus tropicalis
ORGANISM          Xenopus tropicalis
REFERENCE          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
                  Xenopodinae; Xenopus; Silurana.
                  Kremnitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
                  Mardis, E. and Wilson, R.
                  A physical map of the xenopus tropicalis genome
                  Unpublished (2003)
JOURNAL            Contact: Richard K Wilson
                  Genome Sequencing Center
                  Washington University School of Medicine
                  Email: submlsions@wustl.edu
                  Insert Length: 75000 Std Error: 0.00
                  Seg primer: T7 TAATACGACTCACTATAGG
                  Class: BAC ends
                  High quality sequence start: 2
                  High quality sequence stop: 672.

```

FEATURES

```

source
1..824
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-110A2"
/clone_id="ISB1"
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
      library Segment 1"
ORIGIN

```


Query Match	17.2%	Score 183;	DB 9;	Length 824;
Best Local Similarity	59.5%;	Pred. No. 1.7e-37;		
Matches 309;	Conservative 0;	Mismatches 210;	Indels 0;	Gaps 0;

Oy	125	CAGCAAAATCTTGGACATGATATCAATCCCTCCACCTGGGCAAGAAAGACGAAGTCT	184
Db	167	CAGTGAATCTGTACTCTCTGTGGACATCTCAAGCTGGCATCGAACAAGAAAGTCT	226
Oy	185	CTTACAACTATCTCTTGGACCTCGCTGCTCCGACATCTTGGTCTCTTTTCAATGCT	244
Db	227	CTTACAGTACTCTGTGGCTCTTACCATCTCAGACATCTTGGAGCCAGATTTTCATCATTT	286
Oy	245	TTGATGACCTTCCGTGGAGATTTCATCTTGAACATGACAGATGCTCAGTCCCGACGA	304
Db	287	TTGTGGGCTTATCTCTGCAACAGCAATATCTCACCGCAGAGTGGCCAGTACTTAAATCC	346
Oy	305	AGATCATAGAAAGTGTGGAAATTTCTATCATCCATCCACACTCCATATGATTTACTGTACGT	364
Db	347	ATGTGGTCAAGTCTCTTAGTTCCTCTTAAATCATGCAATCTTTTGGGGACAGATATAC	406
Oy	365	TAACCATATGACAGGATATATCGCTGTCCGACCCGGCTCAAGTACACAACGGTCTATACC	424
Db	407	TGACCGTGGACCGTTATATGTGCATTATGTATCCGCTACAAATATGCTCTTCTCTTACC	466
Oy	425	CAGCCCGACCCCGGAAGTCAATGTGAAGTGTATTATCATCACTGCTTCCAGCACATCC	484
Db	467	CAGAGCGAACCCGTGAGATCAATGTGTTGTTCTTCTCATCGTTTGGACCGGTATAC	526
Oy	485	CCATTTACTGTGGGCCCAACATCTGGACGTGAAGACTAACATCAGCACTCTGTGCATCACG	544
Db	527	CTTTCTACTGGTGAAGTATGTTTGGAGGATCCCGGTAACCCAGAAATGCTAGACTTTA	586
Oy	545	TCTCTATCTGATCCATGCTTACACCGTCTATCTGTGCGCTGCTCAATCTTCTTCACT	604
Db	587	TACTCAATGGACCACTGCTTTAATAATATACTTCAATCCCTGCACCAATATTTTGATTA	646
Oy	605	TGAATCAATCATTTGTATCAAAAGCTCAGAGGAATAAGCA	643
Db	647	CCAACTCTGTACTTATCTACAGGCTGAGAGAAAAAATCA	685

FEATURES	REFERENCE
LOCUS CL649540	AUTHORS
DEFINITION CH213-237P13.SP6 CH213 <i>Gasterosteus aculeatus</i> genomic clone	TITLE
ACCESSION CL649540	JOURNAL
VERSION CL649540.1	COMMENT
KEYWORDS GSS.	
SOURCE	
ORGANISM	
	1269 bp DNA linear GSS 06-JUL-2004
	CH213-237P13 3', genomic survey sequence.
	<i>Gasterosteus aculeatus</i> (three spined stickleback)
	<i>Gasterosteus aculeatus</i>
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
	Acanthomorphi; Acanthopterygii; Percomorphi; Gasterosteiformes;
	Gasterosteidae; <i>Gasterosteus</i> .
	1 (bases 1 to 1269)
	Kingstaley,D., Grimwood,J., Dickson,M., Schmutz,J. and Myers,R.M.
	Expressed sequence tags from <i>Gasterosteus aculeatus</i>
	Unpublished (2004)
	Contact: Grimwood, Jane
	Stanford Human Genome Center
	Stanford University School of Medicine
	975 S California Avenue, Palo Alto, CA 94304, USA
	Tel.: 650 320 5917
	Fax: 650 320 5801
	Email: jane@hgc.stanford.edu
	Plate: 237
	Class: BAC ends
	High quality sequence start: 19
	High quality sequence stop: 865.
	Location/Qualifiers
	1..1269

```

/organism="Gasterosteus aculeatus"
/mol_type="genomic DNA"
/strain="Salmon River"
/db_xref="taxon:69293"
/clone="CH213-237F13"
/sex="Mixed"
/cell_type="Blood"
/clone_1b="CH213"
/notes="Vector: pTARAPC2.1; Site 1: EcoRI; The sequence of the clone was established as a mapping and sequencing collaboration at the Stanford Genome Evolution Center, funded by the NIH Centers of Excellence in Genomic Science (CEGS) initiative (http://cegs.stanford.edu). The clone was isolated from the BAC library CHORI-213 built by Pieter de Jong in collaboration with the Stanford Genome Evolution Center (http://www.chori.org/bacpac/). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering\_information.htm)."
```

	Query Match	Best Local Similarity	15.2%;	Score 172.4;	DB 9;	Length 1269;				
	Matches	215;	Conservative	0;	Mismatches	71;	Indels	0;	Gaps	0;
QY	777	CCGCTGCTGTGTGACATCATATGTCACCATTTGCCACATCTAGCCCTTCTGAAACAGAC	836							
DB	42	CGGGCAGCTCTCCACGTGCTCAACCGACCTGGCCATTATCTGCGTTGCTCAACACCGG	101							
QY	837	CATCAACTTTCCTCTACTGCTTCATCAGCAAGCGTTTCGGACCATGACCGCCGAC	896							
DB	102	GGTCAACTTCTCTCTTACTGCTTCATCAGCAAGCGTTTCGGCGGCAATGCGGCCAAGT	161							
QY	897	GTCTCAAGGCTTTCTTCAAGTGCAGAAAGCAACCTGTACATTTCTACACCAATCAACTT	956							
DB	162	CTGGCGGCGCTTGTCTCACTGCTCGGAAGACGCGCGCGCTTCTTACCCAGCCACAACTT	221							
QY	957	TTTCATTAACAAGTAGCCCTGTGATCTTGGCCGGCAAACTCACATGCATCAAGATGTGTT	1016							
DB	222	TTTCATATCAAGAAGACGCCGTGATCTTCAACCGGCCAACTCCACATGTATTAAGATGTGTT	281							
QY	1017	GTACCCGATAGACAAAATGGAAAACCTATTAAAGTATGCCCGGA	1062							
DB	282	GTACCCGATAGACAAAATGGAAAACCGCTGTGATTTCTCTTTGA	327							

FEATURES	
TITLE	JOURNAL COMMENT
REFERENCE	AUTHORS
ORGANISM	SOURCE
KEYWORDS	VERSION
ACCESSION	DEFINITION
LOCUS	LOCUS
RESULT 12	CL137848
	CL137848
	867 bp DNA linear GSS 05-JAN-2004
	ISB1-110C15.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110C15,
	genomic survey sequence.
	CL137848
	CL137848.1 GI:40631483
	GSS.
	Xenopus tropicalis (western clawed frog)
	Xenopus tropicalis
	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
	Xenopodinae; Xenopus; Silurana.
	1 (bases 1 to 867)
	Kremutzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
	Mardis,E. and Wilson,R.
	A physical map of the xenopus tropicalis genome
	Unpublished (2003)
	Contact: Richard K Wilson
	Genome Sequencing Center
	Washington University School of Medicine
	Email: eubm@slonewatson.wustl.edu
	Insert Length: 75000 Std Error: 0.00
	Seq primer: T7 TAATACGACTCATCATTAAGGC
	Class: BAC ends
	High quality sequence start: 11
	High quality sequence stop: 673.
	Location/Qualifiers

source 1..867
 /organism="Xenopus tropicalis"
 /mol_type="genomic DNA"
 /db_xref="taxon:8364"
 /clone="ISB1-110C15"
 /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"

Query Match 14.1%; Score 150; DB 9; Length 867;
 Best Local Similarity 58.1%; Pred. No. 1,1e-28;
 Matches 264; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 125 CAGCAAAATCTTGAAGATGATCATCTCCCTCCAGTGTGCAAGAAAGACAGATCTT 184
 DB 182 CAGTGAACATGCTGACGTGCTGCACTTCAAGCTGCAATCGAAGAAAGATCTT 241
 QY 185 CCTACAACTATCTTGGCACTGCTGCTGCGACATCTTGTCTCTTTTCATAGT 244
 DB 242 CCTACAGTTACCTGTTGGCTTTACATCTCAGACATTTTGAAGCAATTTTCATCTT 301
 QY 245 TTGTGAACTTCTGTTGAAGATTTCACTTGAACAGAGATCTTCAGTCCCGACA 304
 DB 302 TTGTGGGCTTATCTTGAACAGCAATATCTCCAGAGTCCCGAGTACCTTATTC 361
 QY 305 AGATCATGAAGTGTGGAATTTCTATCATCTCACACCTCATATGATTAATGATACCT 364
 DB 362 ATGTGTGAGTGTCTTGAATTTCTCTCTATATCATGATCTATTTGGGTACAGTATAC 421
 QY 365 TAACCATGAGAGATATGCTGCTGCTGCAACCCGCTCAAGTACCAAGCTCTCATAC 424
 DB 422 TGACCGTGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
 QY 425 CAGCCCGCAGCCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 484
 DB 482 CAGAGCAACCCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
 QY 485 CCTATTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
 DB 542 CTTTCTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
 QY 545 TCTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
 DB 602 TACTCAAGTGAACCACTGCTTATATATATCTT 635

RESULT 13
 CNS02EMP 898 bp DNA linear GSS 01-SEP-2000
 LOCUS Tetrarodon nigroviridis genome survey sequence PUC-Or1 end of clone
 DEFINITION 262F08 of library G from Tetrarodon nigroviridis, genomic survey
 sequence.
 accession AL193966
 version AL193966.1 GI:7832072
 keywords GSS; genome survey sequence.
 source Tetrarodon nigroviridis
 organism Tetrarodon nigroviridis

REFERENCE
 AUTHORS Roest Croollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fzames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W., and Weissenbach,J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetrarodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE
 PUBMED 20396633
 REFERENCE 10835645
 AUTHORS Roest Croollius,H., Jalllon,O., Dasilva,C., Ozouf-Costaz,C.,

TITLE
 JOURNAL MEDLINE
 PUBMED 20359837
 REFERENCE 10899143
 AUTHORS 3 (bases 1 to 898)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequenage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetrarodon.
 location/Qualifiers

FEATURES
 source 1..898
 /organism="Tetrarodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="262F08"
 /note="Genoscope sequence ID : C0AG262DC04SPL-end :
 PUC-Or1"

ORIGIN
 Query Match 14.0%; Score 149; DB 9; Length 898;
 Best Local Similarity 56.9%; Pred. No. 2e-28;
 Matches 292; Conservative 0; Mismatches 220; Indels 1; Gaps 1;

QY 121 TTACAGCAAAATCTTGAAGATGATCATCTCCAGTGTGCAAGAAAGACAGAG 180
 DB 370 TTCCAGTGAACATCTGAGCGGCGGTGATACCCGCTGCGCAAGAG 429
 QY 181 TCCCTCAACATATCTCTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB 430 GCACTGATCACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
 QY 241 GTGTTGTGACTTCTGTTGGAAGATTTCACTTGAATGAATGAATGATGATGATGATGAT 300
 DB 490 ATCTGTGGGCTTCTGTTGAGAGCGCATTTTCCACCGGAAGTCCCGCTCTCTC 549
 QY 301 GACAAGATCATGAAGTGTGGAATTTCTGATCATCTCACCTCCATATGATTAAGTACTGTA 360
 DB 550 TTACTCTAGTACGCCCGCCGAGTTGGCGCAACAGGCTTCATCTGATCCAGTTC 609
 QY 361 CCGTTACCATGAGCAGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 610 CCGCTACCGTGAACCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 669
 QY 421 TACCAAGCCCGCAGCCGGAAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 670 TACCCGCGCCGAGCAGAGATCATGCGGAGTTCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 729
 QY 481 ATCCCTATTAAGTGTGCGCAACATCTGCACTGAAGATGATGATGATGATGATGATGATGAT 540
 DB 730 CTGCGCTTCTTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 788
 QY 541 CAGTGTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 789 CCGGTCTCATCTGAGCAGCAGTACATATCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
 QY 601 ATCTGAAGTCAATCTTGTGTATCAAGTCAAG 633
 DB 849 GTCTCAAGCTTTGATCATCAACAGCTGAGG 881

RESULT 14
 LOCUS AL921815 515 bp mRNA linear EST 06-JUL-2004
 DEFINITION AL921815 PUR-Z14Z2 Danio rerio cdna clone 101-F03-2, mRNA sequence.

REFERENCE	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorphia; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
TITLE	1 Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Filames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
JOURNAL	Bistmate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
MEDLINE	Nat. Genet. 25 (2), 235-238 (2000)
PUBMED	20296633
REFERENCE	10835645
AUTHORS	2 Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Filames,C., Fischer,C., Bouneau,L., Billault,Y., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Genome Res. 10 (7), 939-949 (2000)
MEDLINE	20359837
PUBMED	10899143
REFERENCE	3 (bases 1 to 1107)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.
FEATURES	Location/Qualifiers
SOURCE	1..1107
	/organism="Tetraodon nigroviridis"
	/mol_type="genomic DNA"
	/db_xref="taxon:99883"
	/clone="118A07"
	/clone_1id="G"
	/note="Genoscope sequence ID : COBG118AA04LP1-end : T7"
ORIGIN	.
Query Match	10.5% ; Score 111.8 ; DB 9 ; Length 1107 ;
Best Local Similarity	70.6% ; Pred. No. 1.9e-18 ;
Matches 149 ; Conservative	0 ; Mismatches 62 ; Indels 0 ; Gaps 0 ;
Gy	669 GGGAGAACCAACCGCATCTTGTTACACATTACTTCATCTTGGCCACAATTGGGCCCC 728
Db	70 GGGAGAGAGCAGCGGCATGCTGTGGCACATCACTTCGTCTCTGTGCTGTGGCACCC 129
Gy	729 CCGCATATATGATTTCTTTAACCACTCTATAGGGGGCCCATCCAGAACCCTGGCTGGT 788
Db	130 CAGACACGATGATGTCTATCTAACCACTGTACGTCTGTGGTTACCGCGACTGGCGGT 189
Gy	789 GCACATCATGTCGACATTCGCAACATGTAAGCCCTCTGAACAAGACGATCACTTCTT 848
Db	190 CCACCTGCGCTACGACCTGTTCACATGCTGGCCATGCTCAACACGGGCGCTCACTTCTT 249
Gy	849 CCTTACTGCTTCATCAGCAAGCGGTTCCGC 879
Db	250 CTTGTACTGCTTGGTGCAGCAAGCGGTTCCGC 280

Search completed: December 15, 2004, 02:35:03
Job time : 3348.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 18:35:19 / Search time 56.6035 Seconds
(without alignment)
2237.166 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1 MEHTAHLAANSLSWMSFG.....CIKLVLVYDYDKNGKPIKVP 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1872	100.0	353	5	AAU76416 Novel G-P
2	1872	100.0	353	5	ABU74075 Human G-P
3	1872	100.0	353	5	ABB79517 Human che
4	1872	100.0	353	6	ABP98724 Human AXO
5	1872	100.0	353	6	ABR44437 Human G-P
6	1872	100.0	353	6	ABR82499 Human TGR
7	1872	100.0	591	7	ADP70485 Orphan re
8	1868	99.8	353	4	AAU10068 Chemokine
9	1868	99.8	353	7	ADU87769 G-coupled
10	1868	99.8	353	8	AD179325 NOVI1 pro
11	1868	99.8	353	8	AD056004 Human NOV
12	1868	99.8	353	8	AD028954 Human nov
13	1857	99.2	372	4	AAU10067 Chemokine
14	1857	99.2	372	5	AAE18645 Human G-P
15	1857	99.2	372	6	ABR82521 Human G-P
16	1857	99.2	372	7	ADD18023 Human G-P
17	1857	99.2	372	7	ADJ87767 G-coupled
18	1857	99.2	372	8	AD179323 NOVIa pr
19	1852	98.5	372	8	AD056002 Human NOV
20	1844	98.5	353	5	AAE29236 Human NOV
21	1844	98.5	353	6	ABG71163 Human hum
22	1841	98.3	369	6	ABU09571 Human pro
23	1841	98.3	369	6	ABR62522 Human G-P
24	1756	93.8	345	6	ABR82508 Human TGR
25	1756	93.8	345	8	AD028956 Mouse nov

26	1643	87.8	318	7	ADJ87772 G-coupled
27	1643	87.8	318	8	AD179348 NOVIb pr
28	1643	87.8	318	8	AD056027 Human NOV
29	1638	87.5	333	5	AAE17081 Human G-P
30	1638	87.5	333	7	ABW00814 Human GPC
31	1638	87.5	343	5	ABR79519 Human che
32	1638	87.5	385	5	ABR79518 Human che
33	1638	87.5	385	5	ABR95606 Human GPC
34	1626	86.9	333	5	AAE17229 Human thy
35	1471	78.6	287	4	AAU25559 Human G-P
36	1421	75.9	321	7	ADC12696 Human GPC
37	1317	70.4	265	6	ABP81706 Human G-P
38	1317	70.4	265	6	ABR62524 Human G-P
39	912.5	48.7	313	4	AAU25556 Human G-P
40	742	39.6	340	8	AD028936 Mouse nov
41	742	39.6	348	6	ABR82431 Murine 18
42	737	39.4	374	5	ABJ04070 Human G-P
43	737	39.4	374	5	AAW49155 Human G-P
44	737	39.4	374	6	ABR44438 Human G-P
45	737	39.4	374	6	ABR82505 Human TGR

ALIGNMENTS

RESULT 1	AAU76416	standard; protein; 353 AA.
ID	AAU76416	
AC	AAU76416;	
XX		
DT	08-MAY-2002	(first entry)
XX		
DE	Novel G-protein coupled receptor TGR8.	
XX		
KW	G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;	
XX		
KW	Cytotoxic; antiinflammatory; anticancer; fetal brain;	
XX		
KW	central nervous system disease; circulatory organ disorder; cancer;	
XX		
KW	metabolic disease; immunological disease; gastrointestinal disease;	
XX		
KW	gene therapy; transgenic animal; human.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200194582-A1.	
XX		
PD	13-DEC-2001.	
XX		
PF	01-JUN-2001; 2001WO-JP004643.	
XX		
PR	02-JUN-2000; 2000JP-00170446.	
XX		
PR	23-JUN-2000; 2000JP-00194926.	
XX		
PA	(TAKEDA) TAKEDA CHEM IND LTD.	
XX		
PI	Terao Y, Matsui H, Shintani Y;	
XX		
DR	WPI; 2002-164317/21.	
XX		
DR	N-PSDB; ABR15562, ABR15563.	
XX		
PT	Human fetal brain-originated G-protein-coupled receptor protein TGR8 and	
XX		
PT	encoding DNA, for developing drugs to treat e.g. diseases of the central	
XX		
PT	nervous system or circulatory organs, cancer, and metabolic diseases.	
XX		
PS	Claim 1; Fig 2; 102pp; Japanese.	
XX		
CC	The invention describes a human fetal brain-originated G-protein-coupled	
XX		
CC	receptor protein, or its salt. The protein and encoded DNA are useful for	
XX		
CC	developing drugs to treat e.g. diseases of the central nervous system or	
XX		
CC	circulatory organs, cancer, metabolic diseases, immunological diseases	
XX		
CC	and gastrointestinal diseases. The invention also describes creation of a	
XX		
CC	probe for gene therapy and construction of a transgenic animal. This is	
XX		
CC	the amino acid sequence of the novel G-protein coupled receptor TGR8,	
XX		
CC	described in the method of the invention	

XX SQ Sequence 353 AA;

Query Match 100.0%; Score 1872; DB 5; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.8e-200;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWSPGACGIFVPPVYYSLLCLGLPANILTVIILSQLVAROK 60
 DB 1 MEHTAHLAANSLSWSPGACGIFVPPVYYSLLCLGLPANILTVIILSQLVAROK 60

QY 61 SSYNVLLALAAADILVLFPIFVDFLEDFILNMQMPQVPDKITEVLEFSSIHHSIWTV 120
 DB 61 SSYNVLLALAAADILVLFPIFVDFLEDFILNMQMPQVPDKITEVLEFSSIHHSIWTV 120

QY 121 PLTIDRYIAVCHPLKHTVSYSPARTRKIVSVYITCFELTSPYYMWPNIWTEYISTSVH 180
 DB 121 PLTIDRYIAVCHPLKHTVSYSPARTRKIVSVYITCFELTSPYYMWPNIWTEYISTSVH 180

QY 181 HVLIMHCFVTVLVPCGSIFFILNSIIYVKLRKSNFRLRGYSTGKTTALLFTTISIFATL 240
 DB 181 HVLIMHCFVTVLVPCGSIFFILNSIIYVKLRKSNFRLRGYSTGKTTALLFTTISIFATL 240

QY 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300
 DB 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300

QY 301 KAFKCKOKOPVOFTNNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKIVSP 353
 DB 301 KAFKCKOKOPVOFTNNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKIVSP 353

RESULT 2
 ABJ04075 ID ABJ04075 standard; protein; 353 AA.

AC ABJ04075;
 DT 11-OCT-2002 (first entry)
 DE Human G protein coupled receptor hRUP35.
 XX
 KM Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
 KM hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37.
 OS Homo sapiens.
 PN WO200242461-A2.
 PD 30-MAY-2002.
 PF 26-NOV-2001; 2001WO-US044386.
 PR 27-NOV-2000; 2000US-0253404P.
 PR 12-DEC-2000; 2000US-0255366P.
 PR 20-FEB-2001; 2001US-0270266P.
 PR 20-FEB-2001; 2001US-0270286P.
 PR 06-APR-2001; 2001US-0282032P.
 PR 06-APR-2001; 2001US-0282356P.
 PR 06-APR-2001; 2001US-0282358P.
 PR 06-APR-2001; 2001US-0282365P.
 PR 14-MAY-2001; 2001US-0290917P.
 PR 31-JUL-2001; 2001US-0309208P.
 PA (AREN-) ARENA PHARM INC.
 PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
 DR MPI; 2002-56565/60.
 DR N-85DB; ABT04873.
 PT Novel endogenous and non-endogenous versions of G protein-coupled receptor useful for identification of candidate compounds as receptor

PT agonists or antagonists for use as therapeutic agents.
 XX
 PS Claim 29; Page 70-72; 84pp; English.
 XX
 CC The present invention provides the protein and coding sequences of
 CC several human G-protein coupled receptors (GPCRs). These can be used in
 CC the identification of candidate compounds as receptor agonists or inverse
 CC agonists having applicability as therapeutic agents. The present sequence
 CC is a GPCR protein of the invention
 XX

SQ Sequence 353 AA;

Query Match 100.0%; Score 1872; DB 5; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.8e-200;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWSPGACGIFVPPVYYSLLCLGLPANILTVIILSQLVAROK 60
 DB 1 MEHTAHLAANSLSWSPGACGIFVPPVYYSLLCLGLPANILTVIILSQLVAROK 60

QY 61 SSYNVLLALAAADILVLFPIFVDFLEDFILNMQMPQVPDKITEVLEFSSIHHSIWTV 120
 DB 61 SSYNVLLALAAADILVLFPIFVDFLEDFILNMQMPQVPDKITEVLEFSSIHHSIWTV 120

QY 121 PLTIDRYIAVCHPLKHTVSYSPARTRKIVSVYITCFELTSPYYMWPNIWTEYISTSVH 180
 DB 121 PLTIDRYIAVCHPLKHTVSYSPARTRKIVSVYITCFELTSPYYMWPNIWTEYISTSVH 180

QY 181 HVLIMHCFVTVLVPCGSIFFILNSIIYVKLRKSNFRLRGYSTGKTTALLFTTISIFATL 240
 DB 181 HVLIMHCFVTVLVPCGSIFFILNSIIYVKLRKSNFRLRGYSTGKTTALLFTTISIFATL 240

QY 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300
 DB 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFRMAAATL 300

QY 301 KAFKCKOKOPVOFTNNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKIVSP 353
 DB 301 KAFKCKOKOPVOFTNNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKIVSP 353

RESULT 3
 ABB79517 ID ABB79517 standard; protein; 353 AA.

AC ABB79517;
 DT 23-SEP-2002 (first entry)
 DE Human chemokine-like receptor.
 XX
 KM Chemokine-like receptor; G-protein coupled receptor; receptor; human;
 KM HIV infection; cardiovascular disease; asthma;
 KM chronic obstructive pulmonary disease; ashiama;
 KM vasotrophic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;
 KM antiinflammatory; antiallergic; immunomodulator; gene therapy.
 OS Homo sapiens.
 PN WO200248358-A2.
 PD 20-JUN-2002.
 PF 12-DEC-2001; 2001WO-EP014571.
 PR 14-DEC-2000; 2000US-0255150P.
 PR 02-APR-2001; 2001US-0280110P.
 PR 21-JUN-2001; 2001US-0299474P.
 PA (FARB) BAYER AG.
 PI Smolyar A, Zhu Z, Encinas J, Watanabe S, Okiyami H;
 PT

DR MPI: 2002-547858/58.
 DR N-PSDB; ABN64269, ABN64273.
 XX
 XX New isolated polynucleotide encoding a chemokine-like receptor
 PT polypeptide for treating e.g. asthma, myocardial infarction, human
 PT immunodeficiency virus infection, or chronic obstructive pulmonary
 PT disease.
 XX
 XX Claim 25; Fig 2; 114pp; English.
 XX
 CC The present sequence is the protein sequence of a novel human chemokine-
 CC like receptor of 353 amino acids. The chemokine-like receptor has 7
 CC putative transmembrane domains, consistent with the structure of a G-
 CC protein coupled receptor. Its closest human homologue is C-C chemokine
 CC receptor 3. The novel receptor is expressed at low levels in most
 CC tissues. It is expressed at a high level in phytohaemagglutinin-
 CC stimulated CD8+ cells, but in none of the other immune cells tested. It
 CC may act as a receptor of chemottractant molecules on activated
 CC lymphocytes and be involved in cell trafficking and homing to sites of
 CC infection, inflammation or tissue injury. Regulation of activity of the
 CC novel receptor can therefore be used to treat cardiovascular,
 CC immunological and inflammatory diseases, including asthma and chronic
 CC obstructive pulmonary disease (COPD). The receptor may also be a target
 CC for viruses that reside in the nervous system. Regulating the binding of
 CC ligands, e.g. chemottractant molecules or virus particles, to the
 CC receptor can therefore be used to modulate the immune response to inhibit
 CC viral infections, including HIV infection. A claimed method of reducing
 CC activity of the receptor involves contacting a cell with a reagent
 CC (preferably an antibody, antisense oligonucleotide or ribozyme) to a
 CC product (preferably RNA or a polypeptide) encoded by a polynucleotide
 CC encoding the human chemokine-like receptor in vivo or in vitro. A claimed
 CC method of treating a chemokine-like receptor dysfunction related disease
 CC selected from HIV infection, a cardiovascular disorder, asthma or COPD
 CC uses a reagent that modulates a function of the receptor
 CC
 XX
 XX Sequence 353 AA:
 SQ
 Query Match 100.0%; Score 1872; DB 5; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.8e-200;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEHTHAHLAANSSLSWSPGACGAGFVPVYVYSLLCGLPANILVTIILSQLVARROK 60
 DB 1 MEHTHAHLAANSSLSWSPGACGAGFVPVYVYSLLCGLPANILVTIILSQLVARROK 60
 QY 61 SSYNYLLAALAAADILVLFVFDLLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120
 DB 61 SSYNYLLAALAAADILVLFVFDLLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120
 QY 121 PLTIIDRIYAVGHPDKHTVSYPARTRKVIYSVYITCFLTSIPYWMNPNTEDYISTSVH 180
 DB 121 PLTIIDRIYAVGHPDKHTVSYPARTRKVIYSVYITCFLTSIPYWMNPNTEDYISTSVH 180
 QY 181 HVIIMWICFVYVPGSIFFILNSIYYKLRKSNFPLRGYSKGTALIFTTISPATL 240
 DB 181 HVIIMWICFVYVPGSIFFILNSIYYKLRKSNFPLRGYSKGTALIFTTISPATL 240
 QY 241 WAPRIIMILVHYLGAPIONRMVLHMSDIANMLALTAINFLYCPISKRFRMAATL 300
 DB 241 WAPRIIMILVHYLGAPIONRMVLHMSDIANMLALTAINFLYCPISKRFRMAATL 300
 QY 301 KAFKCKOKOPVOYVYTNHNSITSSPMWSPANSHCIKMLVYQYDNKGPIKIVSP 353
 DB 301 KAFKCKOKOPVOYVYTNHNSITSSPMWSPANSHCIKMLVYQYDNKGPIKIVSP 353

RESULT 4

ABP98724 standard; protein; 353 AA.

ID ABP98724 standard; protein; 353 AA.
 AC ABP98724;
 XX
 DT 27-JUN-2003 (first entry)

XX
 DE Human AXOR-57 protein.
 XX
 XX Antibacterial; antifungal; antiviral; antiprotosol; analgesic;
 KM cytoactive; antidiabetic; anorectic; anabolic; antileptomatic; antegonist;
 KM antiparkinsonian; cardiac; hypotensive; hypertensive; neurotropic;
 KM osteopathic; antiangiogenic; antitumor; antiallergic; antiemetic; agonist;
 KM antileptic; neurotropic; tranquilizer; neuroprotective; antidepressant;
 KM vaccine; gene therapy; G-protein coupled receptor; receptor; infection;
 KM pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; stroke;
 KM Parkinson's disease; osteoporosis; angina pectoris; ulcer; allergy;
 KM vomiting; schizophrenia; depression; dementia; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 PN GB365009-A.
 XX
 PD 13-FEB-2002.
 XX
 PF 11-APR-2001; 2001GB-00009018.
 XX
 PR 11-APR-2001; 2001GB-00009018.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Elshourbagy N, Galtu M, Michalovich D, Shabon U;
 XX
 XX MPI: 2003-203569/20.
 DR N-PSDB; ACC44115.
 XX
 PT New G-protein coupled receptor, AXOR 57, for diagnosing and treating
 PT diseases, such as, pain, cancer, diabetes, obesity, anorexia, asthma,
 PT Parkinson's disease, hypotension, hypertension, urinary retention, and
 PT osteoporosis.
 XX
 XX Claim 1; Page 27-28; 32pp; English.
 PS
 CC The invention relates to the isolation of a novel G-protein coupled (7TM)
 CC receptor AXOR 57 (this sequence) or a DNA fragment encoding a polypeptide
 CC having at least 95% identity with AXOR-57 protein. The protein has been
 CC shown to have homology to the human chromosome 16 clone C17B-H1.036A2
 CC (AC008785). The protein and polynucleotides encoding the protein are used
 CC in screening for compounds that stimulate or inhibit the function or
 CC level of the polypeptide or polynucleotides, such as, agonists and
 CC antagonists. The protein, polynucleotides, and antibodies to the protein
 CC are used in diagnostic kits, to diagnose a disease. Polynucleotides
 CC encoding the protein are used for chromosome localization studies, or for
 CC tissue expression studies. The protein and nucleic acids encoding the
 CC protein are used in vaccines for treating diseases such as, bacterial,
 CC fungal, protozoal, and viral infections, pain, cancer, diabetes, obesity,
 CC anorexia, bulimia, asthma, Parkinson's disease, acute heart failure,
 CC hypotension, hypertension, urinary retention, osteoporosis, angina
 CC pectoris, myocardial infection, stroke, ulcer, allergy, benign prostatic
 CC hypertrophy, migraine, vomiting, anxiety, schizophrenia, depression,
 CC delirium, dementia, severe mental retardation, or Huntington's disease
 CC
 XX
 XX Sequence 353 AA:
 SQ
 Query Match 100.0%; Score 1872; DB 6; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.8e-200;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEHTHAHLAANSSLSWSPGACGAGFVPVYVYSLLCGLPANILVTIILSQLVARROK 60
 DB 1 MEHTHAHLAANSSLSWSPGACGAGFVPVYVYSLLCGLPANILVTIILSQLVARROK 60
 QY 61 SSYNYLLAALAAADILVLFVFDLLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120
 DB 61 SSYNYLLAALAAADILVLFVFDLLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120
 QY 121 PLTIIDRIYAVGHPDKHTVSYPARTRKVIYSVYITCFLTSIPYWMNPNTEDYISTSVH 180
 DB 121 PLTIIDRIYAVGHPDKHTVSYPARTRKVIYSVYITCFLTSIPYWMNPNTEDYISTSVH 180

Db 121 PLTIDRYIAVCHPLKHTVSPARTRKIVSVITCFILTSIPYWMNIMTEDYISTSVH 180
Qy 181 HVLIMHCTFTVYLVPCSIFFIINSLIIVYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
Db 181 HVLIMHCTFTVYLVPCSIFFIINSLIIVYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
Qy 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
Db 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
Qy 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKXSP 353
Db 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKXSP 353

RESULT 5
ABR44437
ID ABR44437 standard; protein; 353 AA.

XX ABR44437;
AC ABR44437;
XX 25-JUL-2003 (first entry)
DT 25-JUL-2003 (first entry)
XX Human G protein-coupled receptor #SEQ ID 2.
DE Human, G protein-coupled receptor; anorectic; eating disorder; obesity.
XX Homo sapiens.
OS Homo sapiens.
XX WO2003027142-A1.
PN 03-APR-2003.
XX 19-SEP-2002; 2002WO-JF009626.
PF 21-SEP-2001; 2001JP-00288278.
PR (YAMA) YAMANOUCHI PHARM CO LTD.
XX PA (YAMA) YAMANOUCHI PHARM CO LTD.
XX PI Matsumoto S, Takaaki J, Kurama T, Saito T, Kamohara M, Soga T,
PI Hiayama H;
XX WPI; 2003-333291/31.
DR N-PSDB; ACC71785.
XX New G protein-coupled receptor and encoded gene, useful in screening
PT preventives or remedies for eating disorders or obesity.
XX Claim 1; Page 34-35; 48pp; Japanese.
XX The invention relates to a novel G protein-coupled receptor. The protein
CC and its encoded gene are useful for screening preventives or remedies for
CC eating disorders or obesity. The current sequence represents a G protein
CC coupled receptor sequence
XX
SQ Sequence 353 AA;

Query Match 100.0%; Score 1872; DB 6; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-200;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEHTAHLAANSLSLWMSPGSACGLGFVPVYYSLILCLGHPANILTVIISQVVAROK 60
Db 1 MEHTAHLAANSLSLWMSPGSACGLGFVPVYYSLILCLGHPANILTVIISQVVAROK 60
Qy 61 SSYNVLLAAADIIIVFVVDLEDFILNQMPQVDPKILEVFFSSIHSTIMTV 120
Db 61 SSYNVLLAAADIIIVFVVDLEDFILNQMPQVDPKILEVFFSSIHSTIMTV 120
Qy 121 PLTIDRYIAVCHPLKHTVSPARTRKIVSVITCFILTSIPYWMNIMTEDYISTSVH 180
Db 121 PLTIDRYIAVCHPLKHTVSPARTRKIVSVITCFILTSIPYWMNIMTEDYISTSVH 180

Qy 181 HVLIMHCTFTVYLVPCSIFFIINSLIIVYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
Db 181 HVLIMHCTFTVYLVPCSIFFIINSLIIVYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
Qy 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
Db 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINPFLYCFISKRFRMAAATL 300
Qy 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKXSP 353
Db 301 KAFKCKQKQPVQFYTNHNSITSSPWISPA NSHCIMLVYQYDKNGKPIKXSP 353

RESULT 6
ABB82499
ID ABB82499 standard; protein; 353 AA.

XX ABB82499;
AC ABB82499;
XX 22-JAN-2003 (first entry)
DT 22-JAN-2003 (first entry)
XX Human TGR20 polypeptide.
DE Human TGR20 polypeptide.
XX G-protein coupled receptor; GPCR, antipsoriatic; antiinflammatory;
XX antidiabetic; nootropic; neuroprotective; antianemic; antidiabetic; human;
XX antiparkinsonian; antidiabetic; TGR20; receptor.
XX Homo sapiens.
OS Homo sapiens.
XX WO200277001-A2.
PN 03-OCT-2002.
XX 08-MAR-2002; 2002WO-US007171.
PF 09-MAR-2001; 2001US-00802803.
PR 16-MAR-2001; 2001US-0276649P.
XX (TTLA-) TTLARIK INC.
XX Tian H, Zhao J, Chen J, Cutler G;
XX WPI; 2003-018881/01.
DR N-PSDB; ABV73364.
XX New G-protein coupled receptor polypeptides and polynucleotides useful
PT for identifying compounds for treating a TGR-associated disorder, e.g.
PT psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
PT disease, anemia.
XX Claim 15; Page 63; 87pp; English.

XX The invention relates to G-protein coupled receptor (GPCR) polypeptides
CC and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
CC TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
CC useful for identifying compounds for treating a TGR-associated disorder,
CC such as psoriasis, inflammatory bowel disease, hyperlipidemia,
CC Parkinson's disease, Huntington's disease, anemia, immune and blood
CC disorders, ulcerative colitis, Crohn's disease or spleen enlargement.
CC They are also useful for identifying cells such as kidney, liver,
CC hypothalamus, colon, adipose, or spleen cells, for forensics and
CC paternity determination, diagnosing diseases and examining signal
CC transduction. The present sequence represents a human TGR20 polypeptide
XX

SQ Sequence 353 AA;

Query Match 100.0%; Score 1872; DB 6; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.8e-200;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEHTAHLAANSLSLWMSPGSACGLGFVPVYYSLILCLGHPANILTVIISQVVAROK 60
Db 1 MEHTAHLAANSLSLWMSPGSACGLGFVPVYYSLILCLGHPANILTVIISQVVAROK 60

QY 61 SSNNYLLAALAAADILVLFIVPVDLLEDFILNMQMPQVPDKIIEVLEFSSIHSTIMITY 120
 DB 61 SSNNYLLAALAAADILVLFIVPVDLLEDFILNMQMPQVPDKIIEVLEFSSIHSTIMITY 120
 QY 121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
 DB 121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
 QY 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
 DB 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
 QY 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALTAINPFLYCFISKRPTMAAATL 300
 DB 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALTAINPFLYCFISKRPTMAAATL 300
 QY 301 KAFKCKQKQPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353
 DB 301 KAFKCKQKQPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353

RESULT 7

ID ADF70485 standard; protein; 591 AA.

AC ADF70485;

DT 12-FEB-2004 (first entry)

DE Orphan receptor ligand-related human protein SegID108.

KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
 KM cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KW GFPuv; Enhanced GFP; EGFP; human.

OS Homo sapiens.

PN W02003071272-A1.

PD 28-AUG-2003.

PF 21-FEB-2003; 2003MO-JP001901.

PR 22-FEB-2002; 2002JP-00045728.

PR 23-JUL-2002; 2002JP-00213949.

PR 11-OCT-2002; 2002JP-00298237.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M,

DR N-PSDB; ADF70587.

PT Transformation of cells with a fusion protein of an orphan receptor

PT protein with a fluorescent protein useful for identification of ligands

PS Disclosure; SEQ ID NO 108; 594pp; Japanese.

CC This invention relates to a novel method of identifying ligands to an

CC orphan receptor protein which comprises transforming cells with DNA

CC encoding a fusion protein of the orphan receptor with a fluorescent

CC protein, so that the fusion protein is expressed in the cells (or cell

CC membranes isolated from them) and contacting the cells with the potential

CC ligand to be tested. A suitable fluorescent protein for incorporation in

CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,

CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the

CC identification of ligands binding to an orphan receptor protein.

CC Sequence 591 AA;

Query Match 100.0%; Score 1872; DB 7; Length 591;
 Best Local Similarity 100.0%; Pred. No. 1.2e-199;
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MBHTAHLANSLSNWSPGSAAGLGFVVVYSLLCGLPANILTVIILSQLVAROK 60
 DB 1 MBHTAHLANSLSNWSPGSAAGLGFVVVYSLLCGLPANILTVIILSQLVAROK 60
 QY 61 SSNNYLLAALAAADILVLFIVPVDLLEDFILNMQMPQVPDKIIEVLEFSSIHSTIMITY 120
 DB 61 SSNNYLLAALAAADILVLFIVPVDLLEDFILNMQMPQVPDKIIEVLEFSSIHSTIMITY 120
 QY 121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
 DB 121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
 QY 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
 DB 181 HVLWICHFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
 QY 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALTAINPFLYCFISKRPTMAAATL 300
 DB 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALTAINPFLYCFISKRPTMAAATL 300
 QY 301 KAFKCKQKQPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353
 DB 301 KAFKCKQKQPVQPYTNHNSITSSPMISPANSHCIKMLVYQYDKNGKPIKVP 353

RESULT 8

ID AAU10068 standard; protein; 353 AA.

AC AAU10068;

DT 14-FEB-2002 (first entry)

DE Chemokine receptor family related protein, NOV11.

KW NOV; cytostatic; psoriasis; noctropic; neuroprotectant;

KW cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;

KW haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;

KW differentiation; proliferation; haematopoiesis; wound healing;

KW angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;

KW haemophilia; allergy; Peridred syndrome; skeletal dysplasia;

KW ischemic injury; neuroepithelial disorder; hepatitis; heart failure;

KW chemokine receptor; chromosome 1.

OS Homo sapiens.

PN W0200170978-A2.

PD 27-SEP-2001.

PF 20-MAR-2001; 2001MO-US009093.

PR 20-MAR-2000; 2000US-0190768P.

PR 20-MAR-2000; 2000US-0190835P.

PR 22-MAR-2000; 2000US-0190972P.

PR 22-MAR-2000; 2000US-0191199P.

PR 24-MAR-2000; 2000US-0191477P.

PR 26-MAR-2000; 2000US-0192557P.

PR 28-MAR-2000; 2000US-0192664P.

PR 28-MAR-2000; 2000US-0192665P.

PR 28-MAR-2000; 2000US-0192846P.

PR 29-MAR-2000; 2000US-0192836P.

PR 31-MAR-2000; 2000US-0193843P.

Key Location/Qualifiers
 Peptide 1..47
 FT /label= Signal_peptide
 FT 48..353
 FT /note= "Mature chemokine receptor related protein, NOV11"

CC The invention relates to novel isolated NOVX nucleic acids and encoded
 CC polypeptides. The nucleic acids, polypeptides and antibodies raised
 CC against the polypeptides are useful for preventing or treating diseases
 CC associated with aberrant NOVX expression or activity e.g., cancer,
 CC obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple
 CC sclerosis. The present sequence represents a NOVX polypeptide of the
 CC invention.

CC Sequence 353 AA;

Query Match 99.8%; Score 1868; DB 8; Length 353;
 Best Local Similarity 99.7%; Pred. No. 1,6e-199;

Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSWSPGACGFGFVPPVYYSLLCTGLPANITLVIIISQVARRK 60
 DB 1 MEHTHAHLAANSLSWSPGACGFGFVPPVYYSLLCTGLPANITLVIIISQVARRK 60
 QY 61 SSYNVLLAALADIVLFFIVDFLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 DB 61 SSYNVLLAALADIVLFFIVDFLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 QY 121 PLTIDRYAVCHPLKHTVSPARTKRVIVSYITTCFLTSIPYWWPNIWTEDYISTSVH 180
 DB 121 PLTIDRYAVCHPLKHTVSPARTKRVIVSYITTCFLTSIPYWWPNIWTEDYISTSVH 180
 QY 181 HVLIMHCFYVLPVCSIFFLINSIIYKLRKSNFRLRGVSTGKTALIFTTSTPATL 240
 DB 181 HVLIMHCFYVLPVCSIFFLINSIIYKLRKSNFRLRGVSTGKTALIFTTSTPATL 240
 QY 241 WAPRIIMILYHLGAPIONRWLVHMSDIANMLALNTAINFLPYCFISKRFTMAATL 300
 DB 241 WAPRIIMILYHLGAPIONRWLVHMSDIANMLALNTAINFLPYCFISKRFTMAATL 300
 QY 301 KAEFKCQKQVOPFYTNHNSITSSPMTSPANSHCIKMLVYQYDKNGKPTKVP 353
 DB 301 KAEFKCQKQVOPFYTNHNSITSSPMTSPANSHCIKMLVYQYDKNGKPTKVP 353

RESULT 12

ADO28954 ADO28954 standard; protein; 353 AA.

AC ADO28954;

DT 29-JUL-2004 (first entry)

DE Human novel GPCR PGR3, SEQ ID NO:53.

XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KM transgenic mouse; neurological disorder; adrenal gland disorder;
 KM colon disorder; intestinal disorder; cardiovascular disorder;
 KM muscular disorder; blood disorder; immune disorder; bone disorder;
 KM joint disorder; metabolic disorder; nutritive disorder; cancer;
 KM kidney disorder; liver disorder; lung disorder; breast disorder;
 KM ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KM skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KM thymus disorder; thyroid disorder; antiparkinsonian; antiemetic;
 KM cytosolic; antinflammatory; vasotropic; antiangiogenic; antiarhythmic;
 KM CNS; central nervous system; respiratory; antiatherosclerotic; antidiabetic;
 KM virologic; hepatotropic; antibacterial; antianemic; antiseborrheic;
 KM dermatological; antitumor; antithyroid; antiallergic; anorectic;
 KM immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KM receptor.

XX Homo sapiens.

XX WO2004040000-A2.

XX 13-MAY-2004.

XX 09-SEP-2003; 2003WO-US028226.

PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.

PA (PRIM-) PRIMAL INC.

PI Gallanaris GA, Bergmann JB, Gragerov A, Hohmann J, Li F,

PI Madsen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;

DR WPI: 2004-390329/36.

DR N-PSDB; ADO28955.

PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.

PS Claim 1, SEQ ID NO 53; 542pp; English.

XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 353 AA;

Query Match 99.8%; Score 1868; DB 8; Length 353;
 Best Local Similarity 99.7%; Pred. No. 1,6e-199;

Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSWSPGACGFGFVPPVYYSLLCTGLPANITLVIIISQVARRK 60
 DB 1 MEHTHAHLAANSLSWSPGACGFGFVPPVYYSLLCTGLPANITLVIIISQVARRK 60
 QY 61 SSYNVLLAALADIVLFFIVDFLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 DB 61 SSYNVLLAALADIVLFFIVDFLEDFILNMQMPQVDPKIIIVLEFSSHTSIWTV 120
 QY 121 PLTIDRYAVCHPLKHTVSPARTKRVIVSYITTCFLTSIPYWWPNIWTEDYISTSVH 180
 DB 121 PLTIDRYAVCHPLKHTVSPARTKRVIVSYITTCFLTSIPYWWPNIWTEDYISTSVH 180
 QY 181 HVLIMHCFYVLPVCSIFFLINSIIYKLRKSNFRLRGVSTGKTALIFTTSTPATL 240
 DB 181 HVLIMHCFYVLPVCSIFFLINSIIYKLRKSNFRLRGVSTGKTALIFTTSTPATL 240
 QY 241 WAPRIIMILYHLGAPIONRWLVHMSDIANMLALNTAINFLPYCFISKRFTMAATL 300

KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
 KW transgenic animal; allergy; gene therapy; hepatocytic; anticonvulsant;
 KW neurotropic; neuroprotective; cardiac; immunosuppressive; anorectic;
 KW virucide; receptor.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..47
 FT Domain /label= Signal_peptide
 FT 32..52
 FT /note= "Transmembrane domain"
 FT Protein 48..372
 FT /label= Human_mature_GCREC-6
 FT Domain 74..100
 FT /note= "Transmembrane domain"
 FT Domain 184..208
 FT /note= "Transmembrane domain"
 PN WO200210387-A2.
 PD
 PD 07-FEB-2002.
 PF 25-JUL-2001; 2001WO-US023433.
 PF
 XX
 PR 27-JUL-2000; 2000US-0221478P.
 PR 03-AUG-2000; 2000US-0223268P.
 PR 21-AUG-2000; 2000US-0227054P.
 PR 08-SEP-2000; 2000US-0231121P.
 PR 13-SEP-2000; 2000US-0232243P.
 PR 15-SEP-2000; 2000US-0232691P.
 PR 22-SEP-2000; 2000US-0235146P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 PI Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR,
 PI Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Wala NK, Hatalla AA,
 PI Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L, Graul RC,
 PI Warren BA, Lee EA, Ding L,
 XX
 DR WPI; 2002-188744/24.
 DR N-PSDB; AAD29672.
 PT New human G-protein coupled receptor polypeptide for diagnosis,
 PT prevention and treatment of cell proliferative, neurological,
 PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic
 XX disorders.
 XX
 XX Claim 1; Page 120; 150pp; English.
 CC The invention relates to novel human G-protein coupled receptors (GCREC)
 CC and their encoding polynucleotides. GCREC is useful as an immunogen for
 CC preparing monoclonal and polyclonal antibodies. GCREC is useful for
 CC diagnosing, treating and preventing a cell proliferative disorder (e.g.,
 CC hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,
 CC Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's
 CC disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,
 CC myocardial infarction), gastrointestinal disorder (e.g., gastritis,
 CC cirrhosis, Crohn's disease), an autoimmune/inflammatory disorder (e.g.,
 CC acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,
 CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,
 CC osteoporosis), and viral infections. GCREC is useful in a number of drug
 CC screening techniques, and to analyse the proteome of a tissue or cell
 CC type. GCREC is useful for creating knockin humanised animals or
 CC transgenic animals to model human diseases, in somatic or germline gene
 CC therapy, to generate a transcript image of a tissue or cell type, for
 CC detecting differences in the chromosomal location due to translocation,
 CC inversion, etc., among normal, carrier or affected individuals, and as
 CC hybridization probes for mapping naturally occurring genomic sequences.
 CC GCREC is useful in Southern or northern analysis, dot blot or other
 CC membrane-based technologies, in PCR technologies, in dipstick, pin,
 CC microarray enzyme linked immunosorbent (ELISA)-like assays, and in
 CC microarrays utilising fluids or tissues from patients to detect altered

CC GCREC expression. The present sequence is human GCREC-6
 XX
 XX Sequence 372 AA,
 Query Match 99.2%; Score 1857; DB 5; Length 372;
 Best Local Similarity 100.0%; Pred. No. 3e-198;
 Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MERTHAHLAANSLSWSGSAAGFVPPVYYSLLCLGLPANTLTVIISQVARRQK 60
 DB 1 MERTHAHLAANSLSWSGSAAGFVPPVYYSLLCLGLPANTLTVIISQVARRQK 60
 QY 61 SSVNYLLAAADIIIVFFIVFDLDFILNMQMPVPPKITEVESSIHISWITV 120
 DB 61 SSVNYLLAAADIIIVFFIVFDLDFILNMQMPVPPKITEVESSIHISWITV 120
 QY 121 PLTIIRYIAVCHPLKXHYVSPARTKTVISVYITGFLTSIPYWWPNIWEDYISTSVH 180
 DB 121 PLTIIRYIAVCHPLKXHYVSPARTKTVISVYITGFLTSIPYWWPNIWEDYISTSVH 180
 QY 181 HVLWIHGFVYLVPCSIFFILNSIYVKLRKSNPRLRGYRGTALIFTTISFATL 240
 DB 181 HVLWIHGFVYLVPCSIFFILNSIYVKLRKSNPRLRGYRGTALIFTTISFATL 240
 QY 241 WAPRILMLIYHLYGAPIONRMLVHIMSDIANMLALTAINFPLCYFSKRFRMAAATL 300
 DB 241 WAPRILMLIYHLYGAPIONRMLVHIMSDIANMLALTAINFPLCYFSKRFRMAAATL 300
 QY 301 KAFPKCKQPOVQFTNNPSTSTSPWISPNASHICIKLVYQYDKNGKPIK 350
 DB 301 KAFPKCKQPOVQFTNNPSTSTSPWISPNASHICIKLVYQYDKNGKPIK 350
 RESULT 15
 ABR62521
 ID ABR62521 standard; protein; 372 AA.
 AC ABR62521;
 XX
 XX 06-NOV-2003 (first entry)
 XX
 DE Human G-protein coupled receptor HGPBMY34.
 XX
 XX HGPBMY34; G-protein coupled receptor; GPCR-P14; GPCR-145;
 KW human; neuroprotective; neurotropic; tranquillizer; antimigraine;
 KW neuroleptic; antianemic; antidepressant; anticonvulsant; antiparkinsonian;
 KW cytostatic; cardiac; hypotensive; antidiabetic; analgesic; anorectic;
 KW anti-HIV; antistimulant; osteopathic; uropathic; antidiabetic; antiallergic;
 KW gene therapy.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 16..36
 FT /label= TM1
 FT /note= "Transmembrane domain 1"
 FT 65..87
 FT /label= TM2
 FT /note= "Transmembrane domain 2"
 FT 109..131
 FT /label= TM3
 FT /note= "Transmembrane domain 3"
 FT 148..166
 FT /label= TM4
 FT /note= "Transmembrane domain 4"
 FT 182..208
 FT /label= TM5
 FT /note= "Transmembrane domain 5"
 FT 227..249
 FT /label= TM6
 FT /note= "Transmembrane domain 6"
 FT 269..288
 FT /label= TM7

```

FT      /note= "transmembrane domain 7"
XX      MO2003050256-A2.
PN      19-JUN-2003.
XX      06-DEC-2002; 2002MO-US039290.
XX      06-DEC-2001; 2001US-0338371P.
XX      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX      Feder JN, Gopal S, Mintier GA, Ramanathan CS;
XX      WPI, 2003-577295/54.
DR      N-PSDB; ACF05275.
XX      New nucleic acid molecule encoding a human G-protein coupled receptor,
PT      HGPBMY34, useful for diagnosing, preventing or treating diseases
PT      involving the receptor, for example Parkinson's disease, dementia,
PT      asthma, hypertension or cancer.
XX      Claim 5; Fig 1A-B; 112pp; English.
XX      The present sequence is the protein sequence of human HGPBMY34, a newly
CC      identified G-protein coupled receptor (GPCR) belonging to the group of
CC      Class A GPCRs and showing homology to the P1am model 7 transmembrane
CC      receptor, rhodopsin family. HGPBMY34, also referred to as GPCR-P14
CC      and/or GPCR-145, is highly expressed in brain (especially in the
CC      amygdala, caudate nucleus, corpus callosum, hippocampus, thalamus,
CC      substantia nigra), spinal cord and pituitary, indicating an association
CC      in neurological systems and conditions. It is also expressed in the bone
CC      marrow and testis. The invention provides HGPBMY34 polynucleotides,
CC      polypeptides and antibodies, expression vectors, host cells and antisense
CC      molecules, methods for screening for modulators of HGPBMY34 activity
CC      and/or function, and methods for diagnosing, treating, preventing and
CC      screening for disorders and diseases associated with abnormal HGPBMY34
CC      activity, including: a disorder related to aberrant G-protein coupled
CC      signaling; a disorder related to aberrant cell cycle regulation;
CC      neurological disorders; anxiety; headache; migraine; schizophrenia; manic
CC      depression; delirium; dementia; severe mental retardation and
CC      dyskinesias, such as Huntington's disease or Gilles de la Tourette's
CC      syndrome; Parkinson's disease; brain disorders; spinal cord disorders;
CC      affective disorders; neoplastic disorders; cardiovascular disorders;
CC      acute heart failure; hypotension; hypertension; angina pectoris;
CC      myocardial infarction; an immunological disorder; immune-related
CC      disorders; endocrinal diseases; growth disorders; neuropathic pain;
CC      obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis
CC      ; psychosis; metabolic disorders; pituitary disorders; urinary retention;
CC      ulcers; allergies; or benign prostatic hypertrophy (all claimed)
XX      SQ
XX      Sequence 372 AA:
XX
Query Match      99.2%; Score 1857; DB 6; Length 372;
Best Local Similarity 100.0%; Pred. No. 3e-198;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MEHTHAHLAANSSLSWSPGACGLGFVPPVYVYSLLCGLPANILFTVIIISQLVAROK 60
DB      1 MEHTHAHLAANSSLSWSPGACGLGFVPPVYVYSLLCGLPANILFTVIIISQLVAROK 60
QY      61 SSNYNLLAALAAADILVLFVFPVDFLEDFILNMQMPQVPDKIIIEVLEFSSHTSIWTV 120
DB      61 SSNYNLLAALAAADILVLFVFPVDFLEDFILNMQMPQVPDKIIIEVLEFSSHTSIWTV 120
QY      121 PLTIIDRYIACVCHPLKXYTVSPARTKVIYSVYITCFLTSPYYPMPNMTEDYISTSVH 180
DB      121 PLTIIDRYIACVCHPLKXYTVSPARTKVIYSVYITCFLTSPYYPMPNMTEDYISTSVH 180
QY      181 HVLIMHCFVTYLVPCGIFPLINSIIYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240
DB      181 HVLIMHCFVTYLVPCGIFPLINSIIYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240

```

```

QY      241 WAPRIIMILYHLYGAPIONRMVLVHMSDIANMLALNTAINPFLYCEISKRFRMAAATL 300
DB      241 WAPRIIMILYHLYGAPIONRMVLVHMSDIANMLALNTAINPFLYCEISKRFRMAAATL 300
QY      301 KAFKCKQKQPVQFYTNHNSITSSPMISPANSHCIIMLVYQYDKNGKPIK 350
DB      301 KAFKCKQKQPVQFYTNHNSITSSPMISPANSHCIIMLVYQYDKNGKPIK 350

```

Search completed: December 14, 2004, 20:23:08
Job time : 59.6035 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:25:14 ; Search time 16.4665 Seconds
(without alignments)
1421.692 Million cell updates/sec

Title: US-09-995-225b-16

Sequence: 1 MHTAHALANSLSMWSPG.....CIKMLVYDYDKXGPKIKVSP 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgnt2_6/prodata/1/1aa/5A_COMB.pep.*
2: /cgnt2_6/prodata/1/1aa/5B_COMB.pep.*
3: /cgnt2_6/prodata/1/1aa/6A_COMB.pep.*
4: /cgnt2_6/prodata/1/1aa/6B_COMB.pep.*
5: /cgnt2_6/prodata/1/1aa/6C_COMB.pep.*
6: /cgnt2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	241	12.9	393	1	US-07-629-1041-3
2	238.5	12.7	398	2	US-08-288-663A-1
3	232	12.4	415	4	US-09-545-944-2
4	231.5	12.4	355	1	US-08-012-988A-2
5	231.5	12.4	355	1	US-08-450-933A-5
6	231.5	12.4	355	3	US-08-446-669-5
7	231.5	12.4	355	3	US-09-239-938-1
8	231.5	12.4	355	4	US-09-886-119A-14
9	231.5	12.4	355	4	US-10-039-659A-13
10	231.5	12.4	355	4	US-09-961-068-1
11	231.5	12.4	355	4	US-09-625-573-5
12	231.5	12.4	355	5	US-09-960-547-1
13	231.5	12.4	355	5	PCT-US95-00476-5
14	226.5	12.1	398	2	US-08-288-663A-15
15	223.5	11.9	355	4	US-08-833-752-9
16	223.5	11.9	355	4	US-09-938-719-9
17	223	11.9	352	4	US-09-029-027B-2
18	222.5	11.9	347	4	US-08-118-270-47
19	222.5	11.9	347	5	PCT-US93-08528-47
20	219	11.7	353	5	US-08-118-270-45
21	219	11.7	353	5	PCT-US93-08528-45
22	219	11.7	357	1	US-08-149-093A-4
23	219	11.7	357	2	US-08-911-245-4
24	219	11.7	357	2	US-08-553-058C-4
25	219	11.7	357	2	US-08-514-451A-4
26	219	11.7	367	3	US-09-170-331-4
27	219	11.7	367	3	US-09-510-473-4

28	219	11.7	367	4	US-09-048-916B-4	Sequence 4, App1
29	219	11.7	384	3	US-09-071-434-3	Sequence 3, App1
30	218	11.6	378	4	US-09-045-583-5	Sequence 5, App1
31	218	11.6	378	4	US-09-534-185-5	Sequence 5, App1
32	217	11.6	355	3	US-09-045-583-53	Sequence 53, App1
33	217	11.6	355	4	US-09-534-185-53	Sequence 53, App1
34	217	11.6	367	2	US-08-454-549-2	Sequence 2, App1
35	217	11.6	367	3	US-08-454-552-2	Sequence 2, App1
36	217	11.6	367	3	US-08-147-592A-6	Sequence 6, App1
37	217	11.6	367	3	US-08-889-108-17	Sequence 17, App1
38	217	11.6	367	3	US-08-292-694A-6	Sequence 6, App1
39	217	11.6	367	4	US-09-743-871B-15	Sequence 15, App1
40	217	11.6	367	4	US-08-986-209A-2	Sequence 2, App1
41	217	11.6	367	5	PCT-US94-10358-17	Sequence 17, App1
42	215.5	11.5	367	4	US-08-405-271A-23	Sequence 23, App1
43	215	11.5	367	3	US-08-676-351-2	Sequence 2, App1
44	213.5	11.4	380	3	US-08-676-351-5	Sequence 5, App1
45	213	11.4	337	1	US-08-153-848-46	Sequence 46, App1

ALIGNMENTS

```
RESULT 1
US-07-629-1041-3
; Sequence 3, Application US/076291041
; Patent No. 5288621
; GENERAL INFORMATION:
; APPLICANT: Gershengorn, Marvin C
; APPLICANT: Straub, Richard E
; TITLE OF INVENTION: PITUITARY TRH RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/629,1041
; FILING DATE: 19901214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D - 995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-629-1041-3

Query Match      12.9%; Score 241; DB 1; Length 393;
Best Local Similarity 22.1%; Pred. No. 1.3e-13;
Matches 81; Conservative 75; Mismatches 123; Indels 86; Gaps 15;

QY 19 PSAGCIGG-VPPVYYSLLC-LGUPANILTYIIISQVARKQSSYN-YLLALAADIL 75
DB 16 PQAALVLEQVQVITLIVITIGGIVGIMVAVVNR--TKMRPTPTNCLVSLAVADM 73
QY 76 VLPFFIVFDLLEDFITNNQMPQVPDKI-----TEVLEFSSIHHSITWIVPL 122
```

```

Query Match          12.7%; Score 238.5; DB 2; Length 398;
Best Local Similarity 22.4%; Pred. No. 2,2e-13;
Matches 80; Conservative 70; Mismatches 120; Indels 87; Gaps 13;

Oy      28 VPAVVVYSLLC-IGEPANILTVIIISQVARRKSSYN-YELALAAADILVLPFIVPVD 85
         | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      26 VVILLVLLIICGIGVGINVVLVYMR--TTHMGTPTNCVYLSLAVADMLV----- 75

Oy      86 LLEDFLINMQMPQVPKI-----IEVLEFSIHTSMTWITVPLTIDRYAVCH 132
         | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      76 -----VAAGLPNTDSDISYGSNWGVGCCTCYLYQVIGNASSCSITAFITIRYAIACH 129

Oy      133 PLKHTVSSYPARTKRVISVYITCFNLSIPYMW-----PNWTF-EDYISTSHVHLIWI 186
         | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      130 PIKAQPLCTFSRAKKIIIFVMA--FTSLYCMMLFFLLDINISYKDAIYISCGYKISR 186

Oy      187 HCFEYVLYVPCSEFFILNSIIYKVL-----RRKSN- 215
         | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      187 YSPFLYMDFGVYFVPEMLATVLYGFIARILFNPIPSDKENSKTKWQNDSTHQTNTLN 246

```

[illegible]

Db 212 ----INPIQVTSFLYLLPMTVISVLYLMLRLKOKSLADEGNANIORCKRSVN 267
Qy 228 AIIFTTSLPATLMAPIIMILVLYGAPIQNRW-----LVHIMSDIANMLALNT 278
Db 268 KMLFVLVLFALCMAPFHIDRLFFSF-----VEEMTESLAIAVNLVHVSGV---LFYLS 320
Qy 279 AINFLYCISKGRFRMAAATLKAPKCKQKQVQFTNNHFTSTSPMISPA 330
Db 321 AVNPITLNLRRFOAFQNVISFHK-----QMSQHD-----POLPPA 360

RESULT 4
US-08-012-988A-2
Sequence 2, Application US/08012988A
Patent No. 5652133
GENERAL INFORMATION:
APPLICANT: Murphy, Philip M.
TITLE OF INVENTION: Cloning and Expression of Human
TITLE OF INVENTION: Macrophage Inflammatory Protein-1 alpha (MIP-1
TITLE OF INVENTION: alpha)/RANTES Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/012,988A
FILING DATE: 19930128
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-118
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-012-988A-2

Query Match 12.4%; Score 231.5; DB 1; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

Qy 22 ACGLGFPVNVYVYSLLCGLPANILTVIILSQLVAROKSSNYVLLAALADIVLFFV 81
Db 31 AFGAQLPPL-VSLVFGVIGLVGNILVVLVQYKRLKNTSI-YLNLALISDLLFFLTP 88
Qy 82 F-VDF-LLEDPLINMGPVQPKIIEVLEFSSIHSTIWTVPRTIDRYAVCH---PLKX 136
Db 89 FWDVYKLDKDWFGDM---CKILSGFYTGYSIFILITLTDRIYALVAVPALRA 144
Qy 137 HTVSYPARTKRVIVSVYITCFLTSP-YVWMPNIWTEDYISTSVNH----- 181
Db 145 RVTTFEVIISIIIMALAI---LASMPGLYFSKTQWTFHTHTCSLHPRHSLREMKLFQAL 201
Qy 182 -----VLIWHCFVTVLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTTAII 231
Db 202 KNLPLGLVPLVPLVMIICYT-----GIILKILRRPNE-----KSKAVRLIF 242
Qy 232 TITSIFATLMAPIIMILVLYGAPI-----QNRMLVHIMSDIANMLALNTAIVNPF 285

Db 243 VMIITFPLFTWTPYNLTILSVFODPLFTHCEOSRL-DLAQVTEVIATYHCCVNPV 301
Qy 286 CFSKRRFR 293
Db 302 AFVGERFR 309

RESULT 5
US-08-450-393A-5
Sequence 5, Application US/08450393A
Patent No. 5707815
GENERAL INFORMATION:
APPLICANT: Charo, Israel
APPLICANT: Coughlin, Shaun
TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT
TITLE OF INVENTION: PROTEIN RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Goddard Castro Huddleson & Tatum
STREET: 5 Palo Alto Square
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306-2155
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,393A
FILING DATE: May 25, 1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Coe, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: UCAL-237/02US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-843-5165
TELEFAX: 415-8857-0663
TELEX: 380816COOLEYPA
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
AVTI-SENSE: NO
US-08-450-393A-5

Query Match 12.4%; Score 231.5; DB 1; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

Qy 22 ACGLGFPVNVYVYSLLCGLPANILTVIILSQLVAROKSSNYVLLAALADIVLFFV 81
Db 31 AFGAQLPPL-VSLVFGVIGLVGNILVVLVQYKRLKNTSI-YLNLALISDLLFFLTP 88
Qy 82 F-VDF-LLEDPLINMGPVQPKIIEVLEFSSIHSTIWTVPRTIDRYAVCH---PLKX 136
Db 89 FWDVYKLDKDWFGDM---CKILSGFYTGYSIFILITLTDRIYALVAVPALRA 144
Qy 137 HTVSYPARTKRVIVSVYITCFLTSP-YVWMPNIWTEDYISTSVNH----- 181
Db 145 RVTTFEVIISIIIMALAI---LASMPGLYFSKTQWTFHTHTCSLHPRHSLREMKLFQAL 201
Qy 182 -----VLIWHCFVTVLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTTAII 231
Db 202 KNLPLGLVPLVPLVMIICYT-----GIILKILRRPNE-----KSKAVRLIF 242
Qy 232 TITSIFATLMAPIIMILVLYGAPI-----QNRMLVHIMSDIANMLALNTAIVNPF 285

Db 243 VIMIFLFWTPYNTLILISVFOFLFTHCEQSRHL-DLAVQVTEVIAYTHCCVNPVY 301
QY 286 CFISKRFR 293
Db 302 AFVGERFR 309

RESULT 6

US-08-446-669-5
; Sequence 5, Application US/08446669
; Patent No. 6132987
; GENERAL INFORMATION:
; APPLICANT: Charo, Israel
; APPLICANT: Coughlin, Shaun
; TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOKINE RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESS: Cooley Godward Castro Huddleston & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306-2155
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,669
; FILING DATE: May 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Neeley, Richard
; REGISTRATION NUMBER: 30,092
; REFERENCE/DOCKET NUMBER: UCAL-237/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-843-5000
; TELEFAX: 415-857-0663
; TELEEX: 380816COOLEYPA
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-446-669-5

Query Match 12.4%; Score 231.5; DB 3; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 AAGLGFEVYVYSLICGLPANITVILISQVARRKSSYNVLLAADIIVLFFIV 81
Db 31 AFGAQLLEPL-YSLVFIQVGNILVAVQYKRLKMTSI-YLLNLAISDLFLFTLP 88
QY 82 F-VDF-LLEDFTLNMQVDPDKIEVLEPSSIHSTIWTVPLTIDRYAVCH--PLKY 136
Db 89 FMIDKLDWVFGDAM-----CKILSGFYTGSLSEIFILLTIDRYLAVHAVFALRA 144
QY 137 HTVSPARTKRVISVITCFLTSLP-YMMPNITWEDYSTVSH-----181
Db 145 RTVTFGVITSLIIMALAI--LASWPGLYFSKTQWETHTHCSLHFPESLREWKLFQAL 201
QY 182 -----VLIWHCFYVILVPCSIFFILNSIIVYKLRKSNFRILRGYSTGKTALIF 231
Db 202 KLNFGVLPLVLMVITCYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TITSIFATLWAPRIIMILYHGYAPL-----QNRWLVIHMSDIANMLALNTAINFLY 285

Db 243 VIMIFLFWTPYNTLILISVFOFLFTHCEQSRHL-DLAVQVTEVIAYTHCCVNPVY 301
QY 286 CFISKRFR 293
Db 302 AFVGERFR 309

RESULT 7

US-09-239-938-1
; Sequence 1, Application US/09239938
; Patent No. 6329510
; GENERAL INFORMATION:
; APPLICANT: Qin, Shixin
; APPLICANT: Newman, Walter
; APPLICANT: Kaasam, Nasim
; APPLICANT: Leukosite, Inc.
; TITLE OF INVENTION: ANTI-CCR1 ANTIBODIES AND METHODS OF USE
; FILE REFERENCE: LKS97-13
; CURRENT APPLICATION NUMBER: US/09/239,938
; CURRENT FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-239-938-1

Query Match 12.4%; Score 231.5; DB 3; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 AAGLGFEVYVYSLICGLPANITVILISQVARRKSSYNVLLAADIIVLFFIV 81
Db 31 AFGAQLLEPL-YSLVFIQVGNILVAVQYKRLKMTSI-YLLNLAISDLFLFTLP 88
QY 82 F-VDF-LLEDFTLNMQVDPDKIEVLEPSSIHSTIWTVPLTIDRYAVCH--PLKY 136
Db 89 FMIDKLDWVFGDAM-----CKILSGFYTGSLSEIFILLTIDRYLAVHAVFALRA 144
QY 137 HTVSPARTKRVISVITCFLTSLP-YMMPNITWEDYSTVSH-----181
Db 145 RTVTFGVITSLIIMALAI--LASWPGLYFSKTQWETHTHCSLHFPESLREWKLFQAL 201
QY 182 -----VLIWHCFYVILVPCSIFFILNSIIVYKLRKSNFRILRGYSTGKTALIF 231
Db 202 KLNFGVLPLVLMVITCYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TITSIFATLWAPRIIMILYHGYAPL-----QNRWLVIHMSDIANMLALNTAINFLY 285
Db 243 VIMIFLFWTPYNTLILISVFOFLFTHCEQSRHL-DLAVQVTEVIAYTHCCVNPVY 301
QY 286 CFISKRFR 293
Db 302 AFVGERFR 309

RESULT 8

US-09-886-319A-14
; Sequence 14, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jorn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
; TITLE OF INVENTION: Active Substances

FILE REFERENCE: 50125/014002
CURRENT APPLICATION NUMBER: US/09/886,319A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: US 60/222,081
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: DE 10030149.5
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 355
TYPE: PRT
ORGANISM: Homo sapiens
US-09-886-319A-14

Query Match 12.4%; Score 231.5; DB 4; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGAGFVPPVYVYSLLCGLPANILTVIILSQLVARROKSSYNYLLAALADLVLPFV 81
DB 31 AFGAQLPPL-VSLVFGVIGVGNILVVLVQYKRLKNTSI-YLLNLAIISDLPLFTLP 88
QY 82 F-VDF-LLEDPLNNQMPQVDPKIEVLEFSSIHSTMTVPLTIDRYAVCH---PLKY 136
DB 89 FMIDYKLDQWVFGDM---CKILSGFYTGYSIFPIILFTIDRYAIVHVAVALRA 144
QY 137 HTVSPARTKRVIVSVYITCFLTSP-YVWMPNIMTEDYISTSVH-----181
DB 145 RTVTFGVITSIIMALAI---LASMPGLVFSKTQWEPHTHTCSLHPHESLREWKLFQAL 201
QY 182 -----VLIWHCFYTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILE 231
DB 202 KNLFGVLVPLVMIICYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TTTSIFATLWAPRIIMILVHLVYAP-----ONRWLHMSDIANMLALNTAINFPLY 285
DB 243 VIMIIFPLFWTPYNNILISVFOFLPTECEQSRHL-DLAVQTEVIAVTHCCVNPVIY 301
QY 286 CFISKRF 293
DB 302 AFVGERFR 309

RESULT 9
US-10-039-659A-13
Sequence 13, Application US/10039659A
Patent No. 6723520
GENERAL INFORMATION:
APPLICANT: Wang, Wei
APPLICANT: Gish, Kurt C.
APPLICANT: Schall, Thomas J.
APPLICANT: Vicari, Alain P.
APPLICANT: Zlotnick, Albert
TITLE OF INVENTION: Antibodies that bind chemokine TECK
FILE REFERENCE: DX0589K18 US
CURRENT APPLICATION NUMBER: US/10/039, 659A
CURRENT FILING DATE: 2002-01-03
PRIOR APPLICATION NUMBER: US 08/887,977
PRIOR FILING DATE: 1997-07-03
PRIOR APPLICATION NUMBER: US 60/021,664
PRIOR FILING DATE: 1996-07-05
PRIOR APPLICATION NUMBER: US 60/028,329
PRIOR FILING DATE: 1996-10-11
PRIOR APPLICATION NUMBER: US 60/048,593
PRIOR FILING DATE: 1997-06-04
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 355
TYPE: PRT
ORGANISM: Homo sapiens
US-10-039-659A-13

Query Match 12.4%; Score 231.5; DB 4; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGAGFVPPVYVYSLLCGLPANILTVIILSQLVARROKSSYNYLLAALADLVLPFV 81
DB 31 AFGAQLPPL-VSLVFGVIGVGNILVVLVQYKRLKNTSI-YLLNLAIISDLPLFTLP 88
QY 82 F-VDF-LLEDPLNNQMPQVDPKIEVLEFSSIHSTMTVPLTIDRYAVCH---PLKY 136
DB 89 FMIDYKLDQWVFGDM---CKILSGFYTGYSIFPIILFTIDRYAIVHVAVALRA 144
QY 137 HTVSPARTKRVIVSVYITCFLTSP-YVWMPNIMTEDYISTSVH-----181
DB 145 RTVTFGVITSIIMALAI---LASMPGLVFSKTQWEPHTHTCSLHPHESLREWKLFQAL 201
QY 182 -----VLIWHCFYTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILE 231
DB 202 KNLFGVLVPLVMIICYT-----GIKILRRPNE-----KSKAVRLIF 242
QY 232 TTTSIFATLWAPRIIMILVHLVYAP-----ONRWLHMSDIANMLALNTAINFPLY 285
DB 243 VIMIIFPLFWTPYNNILISVFOFLPTECEQSRHL-DLAVQTEVIAVTHCCVNPVIY 301
QY 286 CFISKRF 293
DB 302 AFVGERFR 309

RESULT 10
US-09-961-068-1
Sequence 1, Application US/09961068
Patent No. 6723570
GENERAL INFORMATION:
APPLICANT: Qiu, Shixin
APPLICANT: Newman, Walter
APPLICANT: Kassem, Nagim
TITLE OF INVENTION: ANTI-CCR1 ANTIBODIES AND METHODS OF USE
FILE REFERENCE: 1855.1048-011
CURRENT APPLICATION NUMBER: US/09/961,068
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 09/239,938
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 355
TYPE: PRT
ORGANISM: Homo sapiens
US-09-961-068-1

Query Match 12.4%; Score 231.5; DB 4; Length 355;
Best Local Similarity 26.3%; Pred. No. 8.2e-13;
Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 ACGAGFVPPVYVYSLLCGLPANILTVIILSQLVARROKSSYNYLLAALADLVLPFV 81
DB 31 AFGAQLPPL-VSLVFGVIGVGNILVVLVQYKRLKNTSI-YLLNLAIISDLPLFTLP 88
QY 82 F-VDF-LLEDPLNNQMPQVDPKIEVLEFSSIHSTMTVPLTIDRYAVCH---PLKY 136
DB 89 FMIDYKLDQWVFGDM---CKILSGFYTGYSIFPIILFTIDRYAIVHVAVALRA 144
QY 137 HTVSPARTKRVIVSVYITCFLTSP-YVWMPNIMTEDYISTSVH-----181
DB 145 RTVTFGVITSIIMALAI---LASMPGLVFSKTQWEPHTHTCSLHPHESLREWKLFQAL 201
QY 182 -----VLIWHCFYTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILE 231
DB 202 KNLFGVLVPLVMIICYT-----GIKILRRPNE-----KSKAVRLIF 242

CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Bertliner & Carson
STREET: 201 N. Figueroa Street, 5th Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00476
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Bertliner, Robert
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 5555-291
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-977-1001
TELEFAX: 310-977-1003
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-00476-5

Query Match	12.4%	Score 231.5	DB 5	Length 355
Best Local Similarity	26.3%	Pred. No. 8.2e-13		
Matches 81, Conservative	57	Mismatches 105	Indels 65	Gaps 13

```

Qy      22  ACGGAGFVVVVYYSLLCTGIPANILFVILISQTVARQKSSVNYVLTAAADIVLFFIV  81
Db      31  AFAGQQLFPEL-YSLFVAVIGLVGNILVVLVAVQYKELKNNKTSI-YLNLAIISDILLFETLP  88
Qy      82  F-VDF-LLEDPIINMQPVDPVKIIEVLFESSIHTSIWTVPLTIDRYAVCH--PLKY  136
Db      89  FMIDYKAKDWMVGDM-----CKILSGFYTGYSBEIPFILLITIDRLALVHAVFLRA  144
Qy      137  HTVSYPARTRKIVSVYITCFELSLP-YVMMPNWITEYISTSVH-----181
Db      145  RYLVFGVITSIILMALAI---LASMPGLVFSKTQWEPHTHTCSLHPHESLRKMLFOAL  201
Qy      182  -----VLIIHCFVTVLVPCSIFFILNLSIIYVKLRKRSNFRNLGSGTGMILF  231
Db      202  KNLPLGLVLEPLVMIICYT-----GIILKLRLRPNE-----KSKAVVILF  242
Qy      232  TITSIPATLMARPIIMILVHLVQAPL-----QNRMLVHMSDIANMLALNTAINFFLY  285
Db      243  VIMIIPELFMTPNMLITILISVFQDLFTHCECOSRHL-DLAVQYEVIAIYTHCCPNVIY  301
Qy      286  CFISKRRR  293
Db      302  AFVGERFR  309

```

RESULT 14
US-08-288-663A-15
; Sequence 15, Application US/08288663A
; Patent No. 5879896
; GENERAL INFORMATION:
; APPLICANT: HINUMA, Shuji
; APPLICANT: HOSOYA, Masaki
; APPLICANT: ONDA, Haruo
; TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
; TITLE OF INVENTION: AND USE

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
00

NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,663A
FILING DATE: 09-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 198309/1993
FILING DATE: 10-AUG-1993
APPLICATION NUMBER: 286986/1993
FILING DATE: 16-NOV-1993
APPLICATION NUMBER: 325215/1993
FILING DATE: 22-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reznick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 44612
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-288-663A-15

Query Match	12.1%	Score 226.5	DB 2	Length 398
Best Local Similarity	22.6%	Pred. No. 2.6e-12		
Matches 83	Conservative 70	Mismatches 15	Indels 89	Gaps 15

```

OY 19 PGSACGJGF-VIPVYVYSLTLC-IGIPANILVYIISOLVARROSSYN-VLATAAADIL 75
Db 16 POVAVALERQVOTLILVAVICGLGIVGINMVLVMMR--TKMRRTANCIIVSLAVADLM 73
OY 76 VLEFPIVFDVFLLEDFILNMQMPQVDPKI-----IEVLESSTHTSITWITVPL 1222
Db 74 VL-----VAAGLPNIITDSIYSGMWYGVGCCLITLYLOVLGINASSCSITAF 119
OY 123 TIDRYIAVCHPLKYHTVSYPARTRKIVSYVYITCELTISIPIYMW-----PNIWT-EDYIS 176
Db 120 TIBRYIA-CHPIAQLCTPSRAKKIIIFVMA--FTSIYCMLEFPLDINISYKDAIV 175
OY 177 TSVHNVLIHCHCTVYIVLPSCSIFPLINSIIYKL-----210
Db 176 ISCGYKISRNYSPYIYMDPGVFYVMPMILTVYIGPIARILFLNPIPSDEKSKTWKN 235
OY 211 ---RRKSNFRLRG-----YSTGKTTALIFTTISIFATLMAP-RIMILVHYGAPI 257
Db 236 DSHQKKNMNLITNRCFSYVSSKKQVTKMLAVVILFALIMNPIYRTLVVANSKLSSPF 255
OY 258 QNEMLVHIMSIDIANNMLALINTALINEFLYCFISKFRITMAAATLKAPFCQKQVOPVPTNH 317
Db 296 QENWPFIL-----FCRICTIYNSAINPVIYINLSQKFR---AARFKLCNCKQKPTKAKANY 347
OY 318 NPSITSS 324
Db 348 SVALNYS 354

```

```

RESULT 15
US-08-833-752-9
; Sequence 9, Application US/08833752
; Patent No. 6448375
; GENERAL INFORMATION:
; APPLICANT: SAMSON, MICHEL
; APPLICANT: PARMENTIER, MARC
; APPLICANT: VASSART, GILBERT
; APPLICANT: LIBERT, FREDERICK
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,752
; FILING DATE: 9-APR-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: linear
; US-08-833-752-9

Query Match      11.9%; Score 223.5; DB 4; Length 355;
Best Local Similarity 26.8%; Pred. No. 4.3e-12;
Matches 79; Conservative 64; Mismatches 113; Indels 39; Gaps 15;

QY      22 ACGLGFPVYVYSSLLCLGLPANILTVIILSQLVARQKSYNYLLALAAADILVLPFIY 81
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      31 AFGAQLPLPL-YSLVFIYGLVGNILVVLVYKRLKNTSI-YLLNLAISDLPLFIPTLP 88
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      82 F-VDF-LLEDFLNNQMPOVPDKIIEVLEFSSIHHSIMITVPLTIDRYIAVCH--PLKY 136
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      89 FWIDYKLDWMVFGDAM---CKIISGFYTGVSIFFIILLTIDRYIAIVHAFALRA 144
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      137 HTVSYPARTKVIIVSYITCFLTSLIP-YMWMNIMWTEYISNSVH--HVLV--W----- 185
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      145 RTVTGVLTSIIIMAI---IASMGILYFSKTQMEFTHTCSLHPHESLREWKLFQAL 201
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      186 -IHCTVIVVPCISIFILINSIIVYKLRRKSNFRLGSGTKTALLFTTISIFATLWAPR 244
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      202 KLNLFGLVL-PLLVNIMICYIGIKILRRPNEK-----KSAVRLIFVIMIIFFLEWIPY 255
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY      245 IIMILYHLYGAPI-----QNRMLVHIMSDIANMLALNTAINFLYCFISKRR 293
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      256 NLTIIISVFQDFLFTHECEQSHL-DLAVQVTEVIAVTHCCVNEVIYAFVGERRR 309
      |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

Search completed: December 14, 2004, 20:28:13
 Job time : 17.4665 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:23:14 ; Search time 49.3994 Seconds

(Without alignment)
2552.337 Million cell updates/sec

Title: US-09-995-225b-16

Sequence: 1 MHTTHAHLAANSLKMSWSPG.....CIRMLVYQYDKNGKPIKVP 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1872	100.0	353	9	US-09-995-225-16
2	1872	100.0	353	10	US-09-995-225-16
3	1872	100.0	353	14	US-10-094-417-2
4	1872	100.0	353	15	US-10-296-294A-5
5	1872	100.0	353	15	US-10-450-590-2
6	1872	100.0	353	16	US-10-779-104-2
7	1868	99.8	353	10	US-09-813-432-22
8	1868	99.8	353	14	US-10-174-364-22
9	1868	99.8	353	15	US-10-246-583-22
10	1868	99.8	353	16	US-10-689-832-22
11	1857	99.2	350	10	US-09-813-432-60
12	1857	99.2	350	14	US-10-174-364-60
13	1857	99.2	350	15	US-10-246-583-60

14	1857	99.2	350	16	US-10-689-832-60	Sequence 60, Appl
15	1857	99.2	372	10	US-09-813-432-20	Sequence 20, Appl
16	1857	99.2	372	14	US-10-219-834-19	Sequence 19, Appl
17	1857	99.2	372	14	US-10-314-076-2	Sequence 2, Appl
18	1857	99.2	372	14	US-10-174-364-20	Sequence 20, Appl
19	1857	99.2	372	15	US-10-333-946-6	Sequence 6, Appl
20	1857	99.2	372	15	US-10-246-583-20	Sequence 20, Appl
21	1857	99.2	372	16	US-10-689-832-20	Sequence 20, Appl
22	1857	99.2	372	17	US-10-712-615-103	Sequence 103, App
23	1853	99.0	350	10	US-09-813-432-59	Sequence 59, Appl
24	1853	99.0	350	14	US-10-174-364-59	Sequence 59, Appl
25	1853	99.0	350	15	US-10-246-583-59	Sequence 59, Appl
26	1853	99.0	350	16	US-10-689-832-59	Sequence 59, Appl
27	1844	98.5	353	14	US-10-012-140-11	Sequence 11, Appl
28	1841	98.3	369	14	US-10-314-076-4	Sequence 4, Appl
29	1756	93.8	345	14	US-10-094-417-20	Sequence 20, Appl
30	1643	87.8	318	14	US-10-174-364-85	Sequence 85, Appl
31	1643	87.8	318	15	US-10-246-583-85	Sequence 85, Appl
32	1638	87.5	333	14	US-10-079-384-26	Sequence 26, Appl
33	1638	87.5	343	15	US-10-450-530-8	Sequence 8, Appl
34	1638	87.5	356	15	US-10-343-650A-22	Sequence 22, Appl
35	1638	87.5	385	15	US-10-791-932-66	Sequence 66, Appl
36	1471	78.6	287	10	US-09-813-432-57	Sequence 57, Appl
37	1426	76.2	272	10	US-09-813-432-57	Sequence 57, Appl
38	1426	76.2	272	14	US-10-174-364-57	Sequence 57, Appl
39	1426	76.2	272	15	US-10-246-583-57	Sequence 57, Appl
40	1426	76.2	272	16	US-10-689-832-57	Sequence 57, Appl
41	1426	76.2	272	16	US-10-689-832-58	Sequence 58, Appl
42	1426	76.2	265	14	US-10-225-567A-587	Sequence 587, App
43	1317	70.4	265	14	US-10-314-076-14	Sequence 14, Appl
44	1317	70.4	265	14	US-10-314-076-14	Sequence 14, Appl
45	912.5	48.7	313	10	US-09-791-932-63	Sequence 63, Appl

ALIGNMENTS

```
RESULT 1
US-09-995-225-16
; Sequence 16, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Zhong T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
```

```

; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US2002013584A1el Sequence
US-09-995-225-16
```

```

Query Match      100.0%; Score 1872; DB 9; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MHTHAHLAANSSLSWSPGACGIFVPVYYSLLCLGLPANILTVIILSOLVAROK 60
Db 1 MHTHAHLAANSSLSWSPGACGIFVPVYYSLLCLGLPANILTVIILSOLVAROK 60
QY 61 SSVNYLLAAADILVLFVFDVFLLEDFTLNMQMPQVPDKIEVLEFSSIHSTIWTY 120
Db 61 SSVNYLLAAADILVLFVFDVFLLEDFTLNMQMPQVPDKIEVLEFSSIHSTIWTY 120
QY 121 PLTIDRYAVCHPLKHTVSPARTKRVISVYITCFITSIPIYWMNMTEDYISTSVH 180
Db 121 PLTIDRYAVCHPLKHTVSPARTKRVISVYITCFITSIPIYWMNMTEDYISTSVH 180
QY 181 HVLIMHCFVTVVPCGIFFLINSIIYKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
Db 181 HVLIMHCFVTVVPCGIFFLINSIIYKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
QY 241 WAPRIIMLYHLGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFTMAATL 300
Db 241 WAPRIIMLYHLGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFTMAATL 300
QY 301 KAFPCOKOPQVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
Db 301 KAFPCOKOPQVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
```

RESULT 2

```

US-09-995-225-16
; Sequence 16, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Hong T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Friede, Catherine
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
```

```

; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9el Sequence
US-09-995-225-16
```

```

Query Match      100.0%; Score 1872; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MHTHAHLAANSSLSWSPGACGIFVPVYYSLLCLGLPANILTVIILSOLVAROK 60
Db 1 MHTHAHLAANSSLSWSPGACGIFVPVYYSLLCLGLPANILTVIILSOLVAROK 60
QY 61 SSVNYLLAAADILVLFVFDVFLLEDFTLNMQMPQVPDKIEVLEFSSIHSTIWTY 120
Db 61 SSVNYLLAAADILVLFVFDVFLLEDFTLNMQMPQVPDKIEVLEFSSIHSTIWTY 120
QY 121 PLTIDRYAVCHPLKHTVSPARTKRVISVYITCFITSIPIYWMNMTEDYISTSVH 180
Db 121 PLTIDRYAVCHPLKHTVSPARTKRVISVYITCFITSIPIYWMNMTEDYISTSVH 180
QY 181 HVLIMHCFVTVVPCGIFFLINSIIYKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
Db 181 HVLIMHCFVTVVPCGIFFLINSIIYKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
QY 241 WAPRIIMLYHLGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFTMAATL 300
Db 241 WAPRIIMLYHLGAPIQNRWLVIHMSDIANMLALNTAINFLYCFISKRFTMAATL 300
QY 301 KAFPCOKOPQVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
Db 301 KAFPCOKOPQVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKVP 353
```

RESULT 3

```

US-10-094-417-2
; Sequence 2, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030045685A1el Receptors
; FILE REFERENCE: 018781-00811005
; CURRENT APPLICATION NUMBER: US/10/094,417
; PRIOR APPLICATION NUMBER: 2002-06-10
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
```

US-10-094-417-2

Query Match	100.0%;	Score 1872;	DB 14;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 2.3e-163;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MEHTHAAHLAANSLSLWMSGSA	CGJGF	PVVVYSL	LLCIGJ	PANILITV	ILSGL	VAROK	60
Dp	1	MEHTHAAHLAANSLSLWMSGSA	CGJGF	PVVVYSL	LLCIGJ	PANILITV	ILSGL	VAROK	60
Qy	61	SSNYVLLAALAAADITLV	FFIVDD	LEBFI	LNMOPV	DKILEV	FESS	IHTS	120
Dp	61	SSNYVLLAALAAADITLV	FFIVDD	LEBFI	LNMOPV	DKILEV	FESS	IHTS	120
Qy	121	PLTIDRYIAVCHPLK	YHTVSP	PARTRK	VLVS	YICFLT	SIPYWM	PMIMED	180
Dp	121	PLTIDRYIAVCHPLK	YHTVSP	PARTRK	VLVS	YICFLT	SIPYWM	PMIMED	180
Qy	181	HULIWHCFVTVL	VP	CSIF	PLN	SIITVY	L	ARKSNFR	240
Dp	181	HULIWHCFVTVL	VP	CSIF	PLN	SIITVY	L	ARKSNFR	240
Qy	241	MABRIMILIVLH	GABIQ	NRMLVH	MSDI	ANMAL	LNTAIN	FEFL	300
Dp	241	MABRIMILIVLH	GABIQ	NRMLVH	MSDI	ANMAL	LNTAIN	FEFL	300
Qy	301	KAFPKCKQKQ	VP	QYTH	NHNS	ITSS	PA	IS	353
Dp	301	KAFPKCKQKQ	VP	QYTH	NHNS	ITSS	PA	IS	353

RESULT 4
US-10-296-294A-5
; Sequence 5, Application US/10296294A

```

1  APPLICANT: TERAQ, Yasuko
2  APPLICANT: MATSUI, Hideki
3  APPLICANT: SHINTANI, Yasuehi
4  TITLE OF INVENTION: No. US20040029224A1e1 G Protein-Coupled Receptor and its DNA
5  FILE REFERENCE: 2734 USOP
6  CURRENT APPLICATION NUMBER: US/10/296, 294A
7  CURRENT FILING DATE: 2001-11-21
8  PRIOR APPLICATION NUMBER: PCT/JP01/04643
9  PRIOR FILING DATE: 2001-06-01
10 PRIOR APPLICATION NUMBER: JP 2000-170446
11 PRIOR FILING DATE: 2000-06-02
12 PRIOR APPLICATION NUMBER: JP 2000-194926
13 PRIOR FILING DATE: 2000-06-23
14 NUMBER OF SEQ ID NOS: 8
15 SEQ ID NO 5
16 LENGTH: 353
17 TYPE: PRT
18 ORGANISM: Human
19 US-10-296-294A-5

```

Query Match	100.0%	Score 1872;	DB 15;	Length 353;
Best Local Similarity	100.0%	Pred. No. 2,3e-153;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MEHTHAHLAANSLSISWSPGSGACGAFPVVVYYSLLCLGHPANILITVYIISQLVARRQK	60	
Db	1	MEHTHAHLAANSLSISWSPGSGACGAFPVVVYYSLLCLGHPANILITVYIISQLVARRQK	60	
Qy	61	SSNYVLLALAADIIVLCFIYFVVDPLLEDPLINMQMPVQPKIIEVLEPSSIHNSITWTV	120	
Db	61	SSNYVLLALAADIIVLCFIYFVVDPLLEDPLINMQMPVQPKIIEVLEPSSIHNSITWTV	120	
Qy	121	PLTIDRIYAVGCHPLKXHTVSYPARTRKIVSVYITCPLTSLIPYMMPINIMEDYISTSVH	180	
Db	121	PLTIDRIYAVGCHPLKXHTVSYPARTRKIVSVYITCPLTSLIPYMMPINIMEDYISTSVH	180	
Qy	181	HVLWICFCFTYVLPDCSIFPLINSIIYVKLRKRSNFRJRGSTGTALAITTISIPATL	240	

Db	181	HVLWHTCFVYLVPCSIFFILNLSIIYKLRKSNFRLRGVSTGKTALIFITISIPATL	240
Qy	241	WARRITMILYHLXGAPIQNRWLVIHMSDIANMLALTAINFLYCIISKRFRTMAATL	300
Db	241	WARRITMILYHLXGAPIQNRWLVIHMSDIANMLALTAINFLYCIISKRFRTMAATL	300
Qy	301	KAFKCKOKPVOFYTNHNFSITSSPWSIPANSHCIKMLVYQYDNGKGRPIRVP	353
Db	301	KAFKCKOKPVOFYTNHNFSITSSPWSIPANSHCIKMLVYQYDNGKGRPIRVP	353

```

RESULT 5
US-10-450-590-2
Sequence 2, Application US/10450590
Publication No. US20040076585A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
FILE REFERENCE: LI0316 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/450,590
CURRENT FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/254,923
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: US 60/280,110
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: US 60/299,474
PRIOR FILING DATE: 2001-06-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 353
TYPE: RNT
ORGANISM: Homo sapiens
US-10-450-590-2

```

```

Query Match 100.0%; Score 1872; DB 15; Length 353;
Best Local Similarity 100.0%; Pred. No. 2,3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSLWSWSPGACGLGFVPVYVYYSLLCLGLPANIILTVIILISQLVARROK 60
DB 1 MEHTHAHLAANSLSLWSWSPGACGLGFVPVYVYYSLLCLGLPANIILTVIILISQLVARROK 60
QY 61 SSNYVTLALAAADILVLFPIVPNDPLLEDFLNNQMPQVPPKILIEVLSESSIHITSIMITV 120
DB 61 SSNYVTLALAAADILVLFPIVPNDPLLEDFLNNQMPQVPPKILIEVLSESSIHITSIMITV 120
QY 121 PLTIDRYAVAGHPKMYHTVSYPARTKRYIVSVITCFPLTSIPYKMPMINWEDYSTSVH 180
DB 121 PLTIDRYAVAGHPKMYHTVSYPARTKRYIVSVITCFPLTSIPYKMPMINWEDYSTSVH 180
QY 181 HVLIMHCFYVLYLPCSGIFFLINSIIYVKLRKRSNFRLDGSGTKTALPLFTITSIFATL 240
DB 181 HVLIMHCFYVLYLPCSGIFFLINSIIYVKLRKRSNFRLDGSGTKTALPLFTITSIFATL 240
QY 241 WAPRIIMLYLHYPGAPIONRWLVHIMSDIANMLALINTAINFELYCFISKRPRTMAAATL 300
DB 241 WAPRIIMLYLHYPGAPIONRWLVHIMSDIANMLALINTAINFELYCFISKRPRTMAAATL 300
QY 301 KAFPRCQKQOPVQFTNNHNFSTTSSPFWISPA NSHCKMLVYQDKNGKPKIXSP 353
DB 301 KAFPRCQKQOPVQFTNNHNFSTTSSPFWISPA NSHCKMLVYQDKNGKPKIXSP 353

RESULT 6
US-10-779-104-2
; Sequence 2, Application US/10779104
; Publication No. US20040161799A1
; GENERAL INFORMATION:
; APPLICANT: Susan Croll-Kalish
; TITLE OF INVENTION: KOR3Jlike-Proteins and Methods of Modulating KOR3J-Mediated Activity

```

US-10-779-104-2
Sequence 2, Application US/10779104
Publication No. US20040161799A1
GENERAL INFORMATION:
APPLICANT: Andrew J. Murphy
APPLICANT: Susan Croll-Kalish
TITLE OF INVENTION: KOR3JL-like-Proteins and Methods of Modulating KOR3JL-Mediated Activity

```
/ FILE REFERENCE: REG 1000A
/ CURRENT APPLICATION NUMBER: US/10/779,104
/ PRIOR FILING DATE: 2004-02-13
/ PRIOR APPLICATION NUMBER: 60/447,447
/ PRIOR FILING DATE: 2003-02-14
/ PRIOR APPLICATION NUMBER: 60/495,577
/ PRIOR FILING DATE: 2003-08-14
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: PaateSeq for Windows Version 4.0
/ SEQ ID NO: 2
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-779-104-2
```

```
Query Match      100.0%; Score 1872; DB 16; Length 353;
Best Local Similarity 100.0%; Pred. No. 2,3e-163;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MEHTAHAAANSSLSWSPGACGFGFVPPVYYSLCLGLPANILTVIILISQVARRK 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MEHTAHAAANSSLSWSPGACGFGFVPPVYYSLCLGLPANILTVIILISQVARRK 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 SSNYTLAAADILVLFVIFVDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 SSNYTLAAADILVLFVIFVDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 PLTIDRYAVCHPKYHTVSYPARTKRYIVSVYITCTLSIPYWMNWTEDYISTSVH 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 PLTIDRYAVCHPKYHTVSYPARTKRYIVSVYITCTLSIPYWMNWTEDYISTSVH 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 HVLIMHCFYTVLPSCIFPLINSIYYKLRKSNFRLRGYSTGKTAILEFTTSIPATL 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 HVLIMHCFYTVLPSCIFPLINSIYYKLRKSNFRLRGYSTGKTAILEFTTSIPATL 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 WAPRIIMILVHLGAPLQNRMLVHMSDIANMLALNTAINFLPYCFISKRFRPTMAATL 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 WAPRIIMILVHLGAPLQNRMLVHMSDIANMLALNTAINFLPYCFISKRFRPTMAATL 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 KAFPKCKQKQPVQFYTNHNFISITSSPMISPAHSHCIKMLVYQYDKNGKPIKVP 353
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 KAFPKCKQKQPVQFYTNHNFISITSSPMISPAHSHCIKMLVYQYDKNGKPIKVP 353
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 7

```
US-09-813-432-22
/ Sequence 22, Application US/09813432
/ Publication No. US20030148485A1
/ GENERAL INFORMATION:
/ APPLICANT: Taupier Jr., Raymond J
/ APPLICANT: Majumder, Kamud
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Verneet, Corinne A. M.
/ TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
/ FILE REFERENCE: 15966-729
/ CURRENT APPLICATION NUMBER: US/09/813,432
/ PRIOR FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
```

```
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/193,843
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 22
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-813-432-22
```

```
Query Match      99.8%; Score 1868; DB 10; Length 353;
Best Local Similarity 99.7%; Pred. No. 5,3e-163;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MEHTAHAAANSSLSWSPGACGFGFVPPVYYSLCLGLPANILTVIILISQVARRK 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MEHTAHAAANSSLSWSPGACGFGFVPPVYYSLCLGLPANILTVIILISQVARRK 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 SSNYTLAAADILVLFVIFVDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 SSNYTLAAADILVLFVIFVDLLEDFILNMQMPQVDPKIEVLEFSSIHSTIWTV 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 PLTIDRYAVCHPKYHTVSYPARTKRYIVSVYITCTLSIPYWMNWTEDYISTSVH 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 PLTIDRYAVCHPKYHTVSYPARTKRYIVSVYITCTLSIPYWMNWTEDYISTSVH 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 HVLIMHCFYTVLPSCIFPLINSIYYKLRKSNFRLRGYSTGKTAILEFTTSIPATL 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 HVLIMHCFYTVLPSCIFPLINSIYYKLRKSNFRLRGYSTGKTAILEFTTSIPATL 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 WAPRIIMILVHLGAPLQNRMLVHMSDIANMLALNTAINFLPYCFISKRFRPTMAATL 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 WAPRIIMILVHLGAPLQNRMLVHMSDIANMLALNTAINFLPYCFISKRFRPTMAATL 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 KAFPKCKQKQPVQFYTNHNFISITSSPMISPAHSHCIKMLVYQYDKNGKPIKVP 353
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 KAFPKCKQKQPVQFYTNHNFISITSSPMISPAHSHCIKMLVYQYDKNGKPIKVP 353
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 8

```
US-10-174-364-22
/ Sequence 22, Application US/10174364
/ Publication No. US20030216308A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson et al.
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2
/ CURRENT APPLICATION NUMBER: US/10/174,364
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
```

```

; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-174-364-22

Query Match          99.8%; Score 1868; DB 14; Length 353;
Best Local Similarity 99.7%; Pred. No. 5,3e-163;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGACGFGVPVYVYSLLCGLPANILTVIILSQLVAROK 60
DB 1 MEHTHAHLAANSSLSWSPGACGFGFPVYVYSLLCGLPANILTVIILSQLVAROK 60
QY 61 SSNNYLLAALAAADILVLFVFDLLEDFILMOMPOVPDKIIEVLEFSSHTSIWTV 120
DB 61 SSNNYLLAALAAADILVLFVFDLLEDFILMOMPOVPDKIIEVLEFSSHTSIWTV 120
QY 121 PLTIDRYIAVCHPLKXHTVSPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
DB 121 PLTIDRYITVCHPLKXHTVSPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
QY 181 HVLWIHCFVTVLPSCIFFLINSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
DB 181 HVLWIHCFVTVLPSCIFFLINSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
QY 241 WAPRIIMLYHLYGAPIQNRWLVIHMSDIANMLALNTAINFPLYCFISGRFTMAATL 300
DB 241 WAPRIIMLYHLYGAPIQNRWLVIHMSDIANMLALNTAINFPLYCFISGRFTMAATL 300
QY 241 WAPRIIMLYHLYGAPIQNRWLVIHMSDIANMLALNTAINFPLYCFISGRFTMAATL 300
DB 241 WAPRIIMLYHLYGAPIQNRWLVIHMSDIANMLALNTAINFPLYCFISGRFTMAATL 300

RESULT 9
US-10-246-583-22
; Sequence 22, Application US/10246583
; Publication No. US20040058862A1
; GENERAL INFORMATION:
; APPLICANT: Majumder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 22
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-246-583-22

Query Match          99.8%; Score 1868; DB 15; Length 353;
Best Local Similarity 99.7%; Pred. No. 5,3e-163;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGACGFGFPVYVYSLLCGLPANILTVIILSQLVAROK 60
DB 1 MEHTHAHLAANSSLSWSPGACGFGFPVYVYSLLCGLPANILTVIILSQLVAROK 60
QY 61 SSNNYLLAALAAADILVLFVFDLLEDFILMOMPOVPDKIIEVLEFSSHTSIWTV 120
DB 61 SSNNYLLAALAAADILVLFVFDLLEDFILMOMPOVPDKIIEVLEFSSHTSIWTV 120
QY 121 PLTIDRYIAVCHPLKXHTVSPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
DB 121 PLTIDRYITVCHPLKXHTVSPARTKRVISVYITCFLSIPYMMNPMTEDYISTSVH 180
QY 181 HVLWIHCFVTVLPSCIFFLINSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
DB 181 HVLWIHCFVTVLPSCIFFLINSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240
QY 241 WAPRIIMLYHLYGAPIQNRWLVIHMSDIANMLALNTAINFPLYCFISGRFTMAATL 300
DB 241 WAPRIIMLYHLYGAPIQNRWLVIHMSDIANMLALNTAINFPLYCFISGRFTMAATL 300
QY 301 KAFKCKOKOPVOPTTNHNSITSSPMISPAHSHCIKMLVYQYDKNGKPIKVP 353
DB 301 KAFKCKOKOPVOPTTNHNSITSSPMISPAHSHCIKMLVYQYDKNGKPIKVP 353

RESULT 10
US-10-689-832-22
; Sequence 22, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majumder, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729D1V1
; CURRENT APPLICATION NUMBER: US/10/689,832
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-832-22
```

Query Match 99.8%; Score 1868; DB 16; Length 353;
Best Local Similarity 99.7%; Pred. No. 5.3e-163;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSLWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIISQLVARROK 60
DB 1 MEHTHAHLAANSLSLWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIISQLVARROK 60
QY 61 SSYNVLLAALADILVLFPIVVDLLEDFILNMOPQVPDKIIIVLFSSIHSTWTV 120
DB 61 SSYNVLLAALADILVLFPIVVDLLEDFILNMOPQVPDKIIIVLFSSIHSTWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSYITCFLTSIPYWMPIWTEEDYSTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSYITCFLTSIPYWMPIWTEEDYSTSVH 180
QY 181 HVLIMHCFYVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISIFATL 240
DB 181 HVLIMHCFYVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISIFATL 240
QY 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINFLFYCFISKRFRMAAATL 300
DB 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINFLFYCFISKRFRMAAATL 300
QY 301 KAFKFCQOPVOFTYNNHFSITSSPWISPAISHCIRKLVYQYDKNGKPIK 353
DB 301 KAFKFCQOPVOFTYNNHFSITSSPWISPAISHCIRKLVYQYDKNGKPIK 353

RESULT 11

US-09-813-432-60

/ Sequence 60, Application US/09813432
/ Publication No. US20030148485A1
/ GENERAL INFORMATION:
/ APPLICANT: Taupier Jr., Raymond J
/ APPLICANT: Matmuder, Kamud
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Verneet, Corine A. M.
/ FILE REFERENCE: 15966-729
/ CURRENT FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: US/09/813,432
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/193,843
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: PatentIn Ver. 2.1
/ LENGTH: 350
/ TYPE: PRT
/ ORGANISM: Homo sapiens

US-09-813-432-60

Query Match 99.2%; Score 1857; DB 10; Length 350;
Best Local Similarity 100.0%; Pred. No. 5.4e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSLWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIISQLVARROK 60
DB 1 MEHTHAHLAANSLSLWSPGSAAGLGFVPPVYVYSLLCGLPANILTVIISQLVARROK 60
QY 61 SSYNVLLAALADILVLFPIVVDLLEDFILNMOPQVPDKIIIVLFSSIHSTWTV 120
DB 61 SSYNVLLAALADILVLFPIVVDLLEDFILNMOPQVPDKIIIVLFSSIHSTWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSYITCFLTSIPYWMPIWTEEDYSTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSPARTKRVIVSYITCFLTSIPYWMPIWTEEDYSTSVH 180
QY 181 HVLIMHCFYVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISIFATL 240
DB 181 HVLIMHCFYVYLVPCSIFFILNSIIYKLRKSNFRLRGYSTGKTALIFTTISIFATL 240
QY 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINFLFYCFISKRFRMAAATL 300
DB 241 WAPRIIMILYHLYGAPIONRWLVHIMSDIANMLALNTAINFLFYCFISKRFRMAAATL 300
QY 301 KAFKFCQOPVOFTYNNHFSITSSPWISPAISHCIRKLVYQYDKNGKPIK 350
DB 301 KAFKFCQOPVOFTYNNHFSITSSPWISPAISHCIRKLVYQYDKNGKPIK 350

RESULT 12

US-10-174-364-60

/ Sequence 60, Application US/10174364
/ Publication No. US20030216308A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson et al.
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2
/ CURRENT FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: US/10/174,364
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/193,843
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 128
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 60
/ LENGTH: 350
/ TYPE: PRT
/ ORGANISM: Homo sapiens

US-10-174-364-60

Query Match 99.2%; Score 1857; DB 14; Length 350;
Best Local Similarity 100.0%; Pred. No. 5.4e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MEHTAHLANSLSWSPGACGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
Db 1 MEHTAHLANSLSWSPGACGLGFVPPVYVYSLLCGLPANILTVIILSQLVARROK 60
Qy 61 SSYNYLALAAADILVLFVFDLFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWITY 120
Db 61 SSYNYLALAAADILVLFVFDLFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWITY 120
Qy 121 PLTIDRYIAVCHPLKXHTVSYPARTRKIVSVYITCFELTSPYMMNPMTEDYISTSVH 180
Db 121 PLTIDRYIAVCHPLKXHTVSYPARTRKIVSVYITCFELTSPYMMNPMTEDYISTSVH 180
Qy 181 HVLIMHCTVTVLPSCSIFPILNSIIVYKLRKSNFRLRGYSTGKTALIFTTTSIFATL 240
Db 181 HVLIMHCTVTVLPSCSIFPILNSIIVYKLRKSNFRLRGYSTGKTALIFTTTSIFATL 240
Qy 241 MAPRIIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISKRPRTMAATL 300
Db 241 MAPRIIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISKRPRTMAATL 300
Qy 301 KAFKCKQKQPVQPYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350
Db 301 KAFKCKQKQPVQPYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350

RESULT 13
US-10-246-583-60
; Sequence 60, Application US/10246583
; Publication No. US2004005862A1
; GENERAL INFORMATION:
; APPLICANT: Majmuder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-246-583-60

Query Match 99.2%; Score 1857; DB 15; Length 350;
Best Local Similarity 100.0%; Pred. No. 5,4e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 61 SSYNYLALAAADILVLFVFDLFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWITY 120
Db 61 SSYNYLALAAADILVLFVFDLFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWITY 120
Qy 121 PLTIDRYIAVCHPLKXHTVSYPARTRKIVSVYITCFELTSPYMMNPMTEDYISTSVH 180
Db 121 PLTIDRYIAVCHPLKXHTVSYPARTRKIVSVYITCFELTSPYMMNPMTEDYISTSVH 180
Qy 181 HVLIMHCTVTVLPSCSIFPILNSIIVYKLRKSNFRLRGYSTGKTALIFTTTSIFATL 240
Db 181 HVLIMHCTVTVLPSCSIFPILNSIIVYKLRKSNFRLRGYSTGKTALIFTTTSIFATL 240
Qy 241 MAPRIIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISKRPRTMAATL 300
Db 241 MAPRIIMILYHLYGAPIONRWLVHMSDIANMLALNTAINFLYCFISKRPRTMAATL 300
Qy 301 KAFKCKQKQPVQPYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350
Db 301 KAFKCKQKQPVQPYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIK 350

RESULT 14
US-10-689-832-60
; Sequence 60, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majmuder, Kamud
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND AMINO ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729DVI1
; CURRENT APPLICATION NUMBER: US/10/689,832
; PRIOR FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-832-60

Query Match 99.2%; Score 1857; DB 16; Length 350;
Best Local Similarity 100.0%; Pred. No. 5,4e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 121 PLTIDRYIAVCHPLKXHTVSPARTRKIVSVYITCFILSIPIYWMNPNTEDYISTSVH 180
QY 181 HVLIMHCFVTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAIIFFITSIPATL 240
Db 181 HVLIMHCFVTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAIIFFITSIPATL 240
QY 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINFPLVCFISKRFRTMAAATL 300
Db 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINFPLVCFISKRFRTMAAATL 300
QY 301 KAFKCOQOPVOQFTYNNHFSITSSPWISPA NSHCIMLVQYDKNKGPXK 350
Db 301 KAFKCOQOPVOQFTYNNHFSITSSPWISPA NSHCIMLVQYDKNKGPXK 350

RESULT 15
US-09-813-432-20
; Sequence 20, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Majmuder, Kamud
; APPLICANT: Spaderna, Steven K
; APPLICANT: Smithson, Glenda
; APPLICANT: Mezes, Peter S
; APPLICANT: Vernet, Corine A. M.
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729
; CURRENT APPLICATION NUMBER: US/09/813,432
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-432-20

Query Match 99.2%; Score 1857; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 5.8e-162;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTAHLAANSLSWMSGSGACGFPVYVYYSLLCGLPANIITVITIIISQVARROK 60
Db 1 MEHTAHLAANSLSWMSGSGACGFPVYVYYSLLCGLPANIITVITIIISQVARROK 60
QY 61 SSYNTLALAAADIIVLFEFIVFDLLEDFILNMQMPQVPDKITIEVLESSSIHTSIWITV 120
Db 61 SSYNTLALAAADIIVLFEFIVFDLLEDFILNMQMPQVPDKITIEVLESSSIHTSIWITV 120

QY 121 PLTIDRYIAVCHPLKXHTVSPARTRKIVSVYITCFILSIPIYWMNPNTEDYISTSVH 180
Db 121 PLTIDRYIAVCHPLKXHTVSPARTRKIVSVYITCFILSIPIYWMNPNTEDYISTSVH 180
QY 181 HVLIMHCFVTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAIIFFITSIPATL 240
Db 181 HVLIMHCFVTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAIIFFITSIPATL 240
QY 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINFPLVCFISKRFRTMAAATL 300
Db 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALNTAINFPLVCFISKRFRTMAAATL 300
QY 301 KAFKCOQOPVOQFTYNNHFSITSSPWISPA NSHCIMLVQYDKNKGPXK 350
Db 301 KAFKCOQOPVOQFTYNNHFSITSSPWISPA NSHCIMLVQYDKNKGPXK 350

Search completed: December 14, 2004, 20:27:31
Job time : 50.3994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:27:39 ; Search time 3552 Seconds

(without alignments)
4699.680 Million cell updates/sec

Title: US-09-995-225b-16

Sequence: 1 MHTTHALAAASLSLWSPG.....CIKMLVQYDKXGKPIKVP 353

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgna2.1/USPTO.spool/US0999225/Runat.14122004.110733.18024/app.query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=xge -MIMATCH=0.1 -LOOFCG=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0999225_@CGN_1_1_2527_@Runat.14122004.110733.18024 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hgt:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_ste:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1872	100.0	1059	6	AX463227 Sequence
2	1872	100.0	1062	6	BD183203 Novel G P
3	1872	100.0	1062	6	BD183204 Novel G P
4	1872	100.0	1062	6	AX498192 Sequence

5	1872	100.0	1062	6	BD103851	BD103851 Novel G P
6	1872	100.0	1062	6	BD103852	BD103852 Novel G P
7	1872	100.0	1167	6	AY635179	AY635179 Homo sapi
8	1868	99.8	1343	6	AX254977	AX254977 Sequence
9	1857	99.2	1119	6	AX254975	AX254975 Sequence
10	1857	99.2	1130	6	AX375235	AX375235 Sequence
11	1844	98.5	1062	6	AX664703	AX664703 Sequence
12	1844	98.5	1062	6	AX664701	AX664701 Sequence
13	1844	98.5	1062	6	AX664702	AX664702 Sequence
14	1643	87.8	957	6	CQ739127	CQ739127 Sequence
15	1641.5	87.7	1788	6	AX375442	AX375442 Sequence
16	1638	87.5	1002	6	AX463231	AX463231 Sequence
17	1638	87.5	1032	6	AX463230	AX463230 Sequence
18	1638	87.5	1070	6	BD144286	BD144286 Novel G-P
19	1638	87.5	1158	6	AB083594	AB083594 Homo sapi
20	1638	87.5	1167	6	AX463232	AX463232 Sequence
21	1638	87.5	1826	6	AC021089	AC021089 Homo sapi
22	1638	87.5	101882	2	AL139235_0	AL139235 Homo sapi
23	1638	87.5	110000	2	AX342465	AX342465 Sequence
24	1631	87.1	1000	6	AC124438	AC124438 Mus muscu
25	1599.5	85.4	183326	10	AC119798	AC119798 Rattus no
26	1598.5	85.4	140885	2	AC121346	AC121346 Rattus no
27	1598.5	85.4	203372	2	AC114097	AC114097 Rattus no
28	1598.5	85.4	230813	2	AX230119	AX230119 Sequence
29	1471	78.6	864	6	AX375239	AX375239 Sequence
30	1421	75.9	963	6	AX657444	AX657444 Sequence
31	1317	70.4	795	6	AX549301	AX549301 Sequence
32	1317	70.4	795	6	HSA03165	HSA03165 Homo sapi
33	1277.5	68.2	1149	5	AY288414	AY288414 Takifugu
34	994	53.1	558	9	AY255545	AY255545 Homo sapi
35	912	48.7	930	6	AX230116	AX230116 Sequence
36	773	41.3	1113	5	AY288413	AY288413 Takifugu
37	737	39.4	945	9	AY255622	AY255622 Homo sapi
38	737	39.4	1122	6	BD185294	BD185294 Novel G P
39	737	39.4	1122	6	BD140885	BD140885 Novel G P
40	737	39.4	1125	6	AX498182	AX498182 Sequence
41	737	39.4	1189	9	AY288421	AY288421 Homo sapi
42	737	39.4	1962	6	AX375239	AX375239 Sequence
43	733	39.2	1125	6	AX664706	AX664706 Sequence
44	733	39.2	1125	6	AX664704	AX664704 Sequence
45	730.5	39.0	951	6	AX657458	AX657458 Sequence

ALIGNMENTS

RESULT 1
AX463227
LOCUS AX463227 1059 bp DNA
DEFINITION Sequence 1 from Patent WO0246358.
ACCESSION AX463227
VERSION AX463227.1 GI:21886199
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Smolyar, A., Zhu, Z., Encina, J., Watanabe, S. and Okigami, H.

AUTHORS Regulation of human chemokine-like receptor

PATENT: WO 0248358-A 1 20-JUN-2002;

JOURNAL Bayer Aktiengesellschaft (DE)

FEATURES
Source 1.1059 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 3.94e-149 Length: 1059
Pred. No.: 1872.00 Matches: 353
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00%

Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-995-225B-16 (1-353) x AX463227 (1-1059)

```

QY 1 MetGluHsThrHsAlaHsIleuAlaHsAsnSerLeuSerTrpTrpSerProGly 20
DB 1 ATGAGGACACGACGCGCCACCTCGACGCCACAGCTGGCTGTGTTGGTCCCCGGC 60
QY 21 SerIaCyGgIyLeuGlyPheValProValValTyTySerLeuLeuLeuCyLeuGly 40
DB 61 TCGGCTGGCGGCTTGGTTCGTCGCCGTGCTACTACAGCCCTTGGCTGTGCTCGGT 120
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
DB 121 TTACGACAAATATCTTGACAGATATATCTCTCCGAGTGGTGGCAGAAAGACGAG 180
QY 61 SerSerTyRasnTyRleuLeuAlaLeuAlaAlaAspIleLeuValIleuPhePhe 80
DB 181 TCCTCTACAACTATCTCTTGGGACCTGGCTGGCCGACATCTTGGCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetProGlnValPro 100
DB 241 GTGTTGTGACCTTCCTGTTGAGAGATTTCATCTTGAAACATGACAGATGCTCAGTCCCC 300
QY 101 AspLysIleIleGlnValLeuGluPheSerSerIleHsThrSerIleTrpIleThrVal 120
DB 301 GACAAAGTATAGAAAGTGTGAAATCTCATCTCCACCTCCATATGATTAAGTCTGTA 360
QY 121 ProLeuThrIleAspArgTyRleAlaValCysHsProLeuLysTyRHisThrValSer 140
DB 361 CCGTTAACCAATTAACAGATATATGCTGCTGACCCGCTCAAGTACACAGGCTCA 420
QY 141 TyRProAlaArgThrArgLysValIleValSerValTyRleHsCysPheLeuThrSer 160
DB 421 TACCCAGCCCGCACCCGGAAAGTCATTTGATGATTAATCACTCGTCTCGACACAC 480
QY 161 IleProTyRTrpTrpProAsnIleTrpHsGluAspTyRleSerThrSerValHs 180
DB 481 ATCCCTATATACGTGGGCGCAACATCTGAGTGAAGACTATACATGACGCTGTGAT 540
QY 181 HisValLeuIleTrpIleHsCysPheThrValTyRLeuValProCysSerIlePhePhe 200
DB 541 CAGCTCTCATCTGAGTATCACTGCTTACCGCTCACTGTCGCTCCATCTTCTTC 600
QY 201 IleLeuAsnSerIleIleValTyRlysLeuArgArgLysSerAsnPheArgLeuArgly 220
DB 601 ATCTTGAACTCATATCTGTGTACAAAGCTCAGAGAGAAAGCAATTTTCGTCGCTGAC 660
QY 221 TyRSerThrGlyLysThrThrAlaIleLeuPheThrIleHsSerIlePheAlaThrLeu 240
DB 661 TACTCCAGGGGAGACACCGCCATCTTGTTCACCATTAACCTCATCTTGGCCACACTT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTyRHisLeuTyRValAlaProIleGlnAsnArg 260
DB 721 TGGGCCCCCGCATATCATATATTTTACCACTCTATGAGGCGCCATCCAGAAACGCG 780
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
DB 781 TGGCGGTACACATCATGTCCGACATTCGCAACATGATGACCTTGTGAACAGCCATC 840
QY 281 AsnPhePheLeuTyRAspPheIleSerLysArgPheArgThrMetAlaAlaIleThrLeu 300
DB 841 AACTTCTCTCTACTGCTTATCATGACAGAGCGGTTCGACACCATAGGACGCGCACGCTC 900
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyRThrAsnHisAsnPheSer 320
DB 901 AAGGCTTCTTCAAGTCCAGAGCAACCTTACAGCTTCAACCAATCATTAATCTTTTCC 960
QY 321 IleHsSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyR 340
DB 961 ATTAACAAGTACCCCTGGATCTCGCCGCAAACTCACATGACATCAAGATGCTGTGTAC 1020

```

QY 341 GlnTyAspLysArgAsnGlyLysProIleLysValSerPro 353
 DB 1021 CAGTATGACAAATAATGAAATACTTATATAAGATATCCCG 1059

RESULT 2
 BD183203 1062 bp DNA linear PAT 17-JUN-2003
 LOCUS Novel G protein-coupled receptor and its DNA.
 DEFINITION BD183203
 ACCESSION BD183203.1 GI:31875403
 VERSION JP 2002345481-A/3.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 1 (bases 1 to 1062)
 Terao, Y., Matsui, H. and Shintani, Y.
 Novel G protein-coupled receptor and its DNA
 Patent: JP 2002345481-A 3 03-DEC-2002;
 TAKEDA CHEMICAL INDUSTRIES LTD

COMMENT
 OS Homo sapiens (human)
 PN JP 2002345481-A/3
 PD 03-DEC-2002
 PE 01-JUN-2001 JP 2001166688
 FT YASUKO TERAO,HIDEKI MATSUI,YASUSHI SHINTANI
 PC C12N15/09,A61K45/00,A61P1/00,A61P3/00,A61P9/00,A61P25/00, PC
 A61P29/00,
 PC A61P35/00,A61P37/00,C07K14/705,C07K16/28,C12N1/15,C12N1/19, PC
 C12N1/21,
 PC C12N5/10,C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC
 G01N33/53,
 PC G01N33/566,C12N15/00,C12N5/00
 CC Novel G protein-coupled receptor and its DNA
 FH Key Location/Qualifiers
 FT source 1..1062
 FT location/Qualifiers
 1..1062 /organism='Homo sapiens (human)'.
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

FEATURES
 source 1..1062
 location/Qualifiers
 1..1062 /organism='Homo sapiens (human)'.
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'

ORIGIN
 Alignment Scores:
 Pred. No.: 3,95e-149 Length: 1062
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-995-225B-16 (1-353) x BD183203 (1-1062)

```

QY 1 MetGluHsThrHsAlaHsIleuAlaHsAsnSerLeuSerTrpTrpSerProGly 20
DB 1 ATGAGGACACGACGCGCCACCTCGACGCCACAGCTGGCTGTGTTGGTCCCCGGC 60
QY 21 SerIaCyGgIyLeuGlyPheValProValValTyTySerLeuLeuLeuCyLeuGly 40
DB 61 TCGGCTGGCGGCTTGGTTCGTCGCCGTGCTACTACAGCCCTTGGCTGTGCTCGGT 120
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
DB 121 TTACGACAAATATCTTGACAGATATATCTCTCCAGTGGTGGCAGAAAGACGAG 180
QY 61 SerSerTyRasnTyRleuLeuAlaLeuAlaAlaAspIleLeuValIleuPhePhe 80
DB 181 TCCTCTACAACTATCTCTTGGGACCTGGCTGGCCGACATCTTGTCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetProGlnValPro 100
DB 241 GTGTTGTGACCTTCCTGTTGAGAGATTTCATCTTGAAACATGACAGATGCTTCCGCC 300

```

QY 101 AspLysIleIleGluValIleuGluPheSerSerIleThrSerIleThrVal 120
 DB 301 GACAAAGTCTAAGAGGCTGGAATTCATCCATCCACACCTCCATATGATTAACGTA 360
 QY 121 ProLeuThrIleAspArgTrIleAlaValCySHisProLeuThrIleThrValSer 140
 DB 361 CGCTTAACATTAACAGGTATATCGCTGTGCGACCCGCTCAAGTACCAACAGGTCTCA 420
 QY 141 TyrProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCACCCGGAAAGCATTTGTAAGTTTACATCACTGCTTCGACACAGC 480
 QY 161 IleProDyTyrTrpTrpProAsnIleTrpThrGluAspTrpIleSerThrSerValHis 180
 DB 481 ATCCCTATTATACGTGTGGCCCAACATCTGAGACGAACTACATCAGACCTCTGTGCAT 540
 QY 181 HisValIleuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CACGTCCTATCTGGATCCACCTGCTTCAACGCTTACCTGCTGCTGCTCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 601 ATCTTAACATCAATCATTTGTGTACAAAGCTCAGAGAGAAAGCAATTTTGGTCCGTGCG 660
 QY 221 TyrSerThrGlyTyrThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCCAGCGGGAAGACACACCGCATCTGTGTCAACCTTCACTTGTGACACCTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCGCCGATCATCATGATTTCTTTACCACTCTCATGGGGCGCCATCCAGAACCGC 780
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 781 TGGCTGTACATCATGTCTGCAATTCGCAATGCAATGCTTACCTTGAACACAGCATC 840
 QY 281 AsnPhePheLeuTyrCysPheIleSerValArgPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACTCTTCTCTCACTGCTTCATCAGCAAGCGGTTCGCAACATGAGCGCGCACGCTC 900
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGGCTTTCTTCAAGTGCCAGAAAGCAACCTGTACAGTTCACATCACTTAATCTTTCC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 961 ATTAACAAGTACCCCTGATCTCGCGGCAAACTCACATGCAATCAAGATGTGTGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1021 CAGTATGACAAATAATGAAACCTATTAAGATATCCCGC 1059

RESULT 3
 BD183204 1062 bp DNA linear PART 17-JUN-2003
 LOCUS BD183204
 DEFINITION Novel G protein-coupled receptor and its DNA.
 ACCESSION BD183204.1 GI:31875404
 VERSION JP 2002345481-A/4.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 1062)
 Terao, Y., Matsui, H. and Shintani, Y.
 Novel G protein-coupled receptor and its DNA
 Patent: JP 2002345481-A 4 03-DEC-2002;
 TAKEDA CHEMICAL INDUSTRIES LTD
 OS Homo sapiens (human)
 PN JP 2002345481-A/4
 PD 03-DEC-2002
 PF 01-JUN-2001 JP 200116668
 PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

PC C12N15/09, A61K45/00, A61P1/00, A61P3/00, A61P9/00, A61P25/00, PC
 A61P29/00,
 PC A61P35/00, A61P37/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, PC
 C12N1/21,
 PC C12N5/10, C12P21/02, C12Q1/58, G01N33/15, G01N33/50, G01N33/53, PC
 G01N33/53,
 PC G01N33/566, C12N15/00, C12N5/00
 CC Novel G protein-coupled receptor and its DNA
 FH Key Location/Qualifiers
 FT source 1. 1062
 FT /organism="Homo sapiens (human)".
 FEATURES
 source Location/Qualifiers
 1. 1062
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 3,95e-149 Length: 1062
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-995-225B-16 (1-353) x BD183204 (1-1062)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGAGACACACGACGCCACCCACCTCGACGACCAAGCTCGCTGTGGTGTCCCGCGC 60
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
 DB 61 TCGCGCTGGCGCTTGGGTTTCCGCGCGTGTCTTACAGCTCTTGTGCTGCTCGGT 120
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlyLeuValAlaArgArgGlnLys 60
 DB 121 TTACACGAAATATCTTGACAGATCATCTCTCCAGCTGGTGGCAGAAACAGAAAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
 DB 181 TCCTCTTACACATATCTCTTGGCCTCGCTGCGCGGACATCTGGATCTCTTTTCATA 240
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 241 GTGTTGTGGACTTCTGTTGGAAGATTTCATCTTGAACATGAGATGCTCAGTCCCGC 300
 QY 101 AspLysIleIleGluValIleuGluPheSerSerIleHisThrSerIleThrValSer 120
 DB 301 GACAAAGTCTAAGAGGCTGGAATTCATCCATCCACACCTCCATATGATTAACGTA 360
 QY 121 ProLeuThrIleAspArgTrIleAlaValCySHisProLeuThrIleThrValSer 140
 DB 361 CGCTTAACATTAACAGGTATATCGCTGTGCGACCCGCTCAAGTACCAACAGGTCTCA 420
 QY 141 TyrProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCACCCGGAAAGCATTTGTAAGTTTACATCACTGCTTCGACACAGC 480
 QY 161 IleProDyTyrTrpTrpProAsnIleTrpThrGluAspTrpIleSerThrSerValHis 180
 DB 481 ATCCCTATTATACGTGTGGCCCAACATCTGAGACGAACTACATCAGACCTCTGTGCAT 540
 QY 181 HisValIleuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CACGTCCTATCTGGATCCACCTGCTTCAACGCTTACCTGCTGCTGCTCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 601 ATCTTAACATCAATCATTTGTGTACAAAGCTCAGAGAGAAAGCAATTTTGGTCCGTGCG 660
 QY 221 TyrSerThrGlyTyrThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240

Db		661 TACTCCACGGGAGAACCCACCGCACTTGTTCACCATTAACCCATCTTTGCCACACTT	720
OY	241 TPRLAProArgIleIleMetIleLeuYrHileuYrGIYAlAProIIegInasnaYg	260	
Db	721 TGGGCCCCCGGATCATTCACTGAATTTCTTTACCACTTAAGGAGCGCCATCCAGAACC	780	
OY	261 TripleValHisIleMetSerAspIleAlaasemtleualaleuleubanthralaje	280	
Db	791 TGGCGTGGCACATCATGTCCGACATTTGCCACAATGTACGCCCTTCTGAAACAGCCATC	840	
OY	281 AsnPherPeLeuYrCyaspheIleaserIysalngphedagThMeValalaalathreu	300	
Db	841 AACTCTTCTCTACTGCTTCAATCAGCAAGCGGTTCCGACCATGTGACGCCCAAGCTC	900	
OY	301 LysalaphelyleCYeGlnlyesGlnProValGlnPheYrThAsnHisasnphseer	320	
Db	901 AAGGCTTCTTCAAGTGCAGAGAACCTGTACAGTTCTACACCAATCATDACTTTTCC	960	
OY	321 IleThrSerSerProTrpIleserProAlaanSerHisCYsileYmeleuValYr	340	
Db	961 ATAAcAAGTAGACCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATCTGTGTAC	1020	
OY	341 GlnTYrAspLYsaenGLyEsrProIleystValserPro	353	
Db	1021 CAGTATGACAAAATGCAAAACCTTATAAAGTATCCCCG	1059	

RESULT 4			
AX498192			
LOCUS			
AX498192	1062 bp	DNA	
DEFINITION	Sequence		linear
AX498192	15 from Patent WO0242461.		PAT 26-SEP-2002
ACCESSION			

VERSION	AX498192.1	GI:23343117
KEYWORDS		

SOURCE ORGANISM	synthetic construct synthetic construct artificial sequences
1	

REFERENCE	AUTHORS	TITLE
1	Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.	Endogenous and non-endogenous sources of human capital

JOURNAL OF CELLULAR PHYSIOLOGY 189:111-120 (2002)
 receptors
 Patent: WO 0242461-A 15 30-MAY-2002;
 Arona Pharmaceuticals, Inc.

Source	Location/Qualifiers	Features
1	Atlanta	Pharmaceuticals, Inc. (US)

```

source
1..1062
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Novel Sequence"

```

ORIGIN

Alignment Scores:

Pred. No.:	3.95e-149	Length:	1062
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 6 Gaps: US-09-995-225B-16 (1-353) X AX498192 (1-1062)

1 MEGTHISTRHISALHISLEVALAALAASNSERSERLEUSERTPTPTPSERPROGLY 20

1 ATGAGAGCACACGACGCCCACTTCGACGACCAACAGCTCGCTGTCTTGTTGGTGTCCCCCGGC 60

61 TGGGCTGGGGCTTGGTTTCGTGCCCGAGGTCTACTACAGAGCTCTTGCTGTGCTGGGT 120

41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaIargArgIleLys 60

121 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGCGAAGAAGACAGAAG 180

QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaIleLeuValLeuPhePheIle 80
|||
Db 181 TCCTCTCAACTATCTCTTGCACTCGTGTGCGGACATCTTGTCCTTTTTCATA 241

QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100

[illegible]

D_b 301 GACAGATCATGAGTCTGGAATTCTCATCCATCCACACCTTCATATGATTACTGTA 360

Qy 122 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuIysTyrHisThrValSer 140

141 TyrProAlaArgThrGlyValIleValSerValTyrIleThrCysPheLeuThrSer 160
161 ThrValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 180
181 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 200
201 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 220
221 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 240
241 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 260
261 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 280
281 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 300
301 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 320
321 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 340
341 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 360
361 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 380
381 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 400
401 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 420
421 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 440
441 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 460
461 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 480
481 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 500
501 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 520
521 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 540
541 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 560
561 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 580
581 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 600
601 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 620
621 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 640
641 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 660
661 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 680
681 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 700
701 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 720
721 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 740
741 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 760
761 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 780
781 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 800
801 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 820
821 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 840
841 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 860
861 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 880
881 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 900
901 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 920
921 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 940
941 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 960
961 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 980
981 GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1000

Db 421 TACCAAGCCGACCCGGAAGTCATTGTAGTGTTCATCATCACTGCTTCTGACACAGC 480

Qy 161 IleProTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180

181 HisValleuileTrrpleHieCyospherhrvalTyrlenuValProCySerilephe 200
|||||
QyencltsgtssccchrcvAlrlEGsAClGAGAGATACATCAGCACCTCTGTGCAT 540

Db 541 CAGCTCCTATCTGAGATCCACAGCTTACACCGTCTACCTGTGCGCCCTGCTCCACTTCTTC 600

201 IleLeuAnsSerIleIleValTyrLysLeuArgArgLysSerAsnBpheArgLeuArgGly 220

QY
221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 660

661 TACTCCAGGGGAAGACCAACGCCCATTTGTTACCACTTAACCTCCATCTTGGCCACTT 720

241 TPALAProArgIleIleMetIleLeuTyrHisLeuYrGlyAlaProIleGlnAsnArg 260

261 TrrleuValHisileMetSerAspIlealaenMetLeualaLeuleuanThrIalile 280
incncacccicmccccccatccagAACCCG 780

781 TGGCTGGTGCACATCATGTCCGACATTGCCACATGTAAGCCCTTCTGAACACAGCCATC 840

281 Asnphheleuryrcyashel leserlysarqphneqrqThrmwAlaAlaAlaThleu 300
841 AACTTCTCCTCACTGCTTCATAGCAAGAGTTTCCCAACCATTTCCACCTTTTCTTTT

301 LysAlaIaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnIleSer 320
1000 LysAlaIaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnIleSer 500

901 AAGCTTCTCAAGTGCAGAACCACTGTACAGTTCTACACCAATCATTAATTTC 960

321 IleInSerSerProTrrPleSerProAlaAsnSerHisCysAlleIysMetLeuValTyr 340
961 ATACAGTAGCCCTCGATCTGCGCGGGAACCTGACATGGTATCAATATCTCTCTTC 1000

341 GlnTyrAspIysAsnGlyLysProIleLeuValSerPro 353

1021 CAGTATGACAAAAATGGAAAACTATAAAGTATCCCG 1059

	DNA	14mer	PAT 37 ATG 0000
BD103851	1062 bp		
OCTUS			
BD103851			
RESULT 5			

REFINITION	Novel G protein-coupled receptor and its DNA.	FBI 2/-AUG-2002
ACCESSION	BMI03851	
DESCRIPTION	Novel G protein-coupled receptor and its DNA.	
ACCESSION	BMI03851	
DESCRIPTION	Novel G protein-coupled receptor and its DNA.	

KEYWORDS
DATE
DOI: 10.1002/ajpa.20000
MO 0194582-A/3.
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo

REFERENCE
AUTHORS
1 (bases 1 to 1062)
Terao, Y., Matsui, H. and Shinicani, Y.

TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: WO 0194582-A 3 13-DEC-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

COMMENT OS Homo sapiens (human)
PN MO 0194582-A/3
PD 13-DEC-2001
PR 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP 00P 170446, 23-JUN-2000 JP 00P 194926 PI
YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00, A61P29/00,
PC A61P9/00, A61P35/00, A61P3/00, A61P37/00, A61P1/00, G01N33/566, PC G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
FT Location/Qualifiers
FT /organism='Homo sapiens (human)'.
/db_xref='taxon:9606'

ORIGIN

Alignment Scores:
Pred. No.: 3,95e-149 Length: 1062
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x BD103851 (1-1062)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20
DB 1 ATGGACACACGACGCCACCTCGACGCCAACAGCTCGCTGCTGGTGCTCCCGGC 60
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB 61 TGGGCTGGGGCTGGGCTTCGTCGCCCGCTGCTACTACAGCTCTTGGCTCTCGGT 120
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGln 60
DB 121 TTACACGACAAATATCTTGACAGATCATCTCTCCAGCTGGTGCAGAACAGAG 180
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
DB 181 TCCTCTACCAACTATCTTGGACCTCGCTGCCGACATCTTGGCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 241 GCTTTGTGACACTCTCTTGGAGAAATTCATTTGAACATGACAGATGCTCAGTCCC 300
QY 101 AspLeuIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
DB 301 GACAAATCATAGAAATGCTGGAATCTCATCATCCACACCTCCATATGATTAAGT 360
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLeuTyrHisThrValSer 140
DB 361 CCGTTAACATTGACAGGTATATGCTGCTGCCACCGCTCAAGATACACAGGCTCA 420
QY 141 TyrProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 421 TACCCAGCCCGCACCCGGAAGTCATTTGAAGTGTATACATCACTCGCTCTCGACAC 480
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
DB 481 ATCCCTATTATGAGTGGCCCAACATCTGAGTGAAGACTATACAGACCTCTGTCAT 540
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200

DB 541 CAGCTCTCATCTGGATCCACTGCTTACACCGCTTACTGCTGCTCTTCTTCTTC 600
QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgValSerAsnPheArgLeuArgGly 220
DB 601 ATCTGACTCATCATCTTGTATGACAGTCAAGAGAGCAATTTTGTCTCCGTGGC 660
QY 221 TyrSerThrGlyValThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 661 TACTCCAGCGGGAAGACACCGCATCTTGTTTCAACATTACCTCTTGTGCAACACT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValAspProIleGlnAsnArg 260
DB 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTTATGGGCGCCCTCCAGAACCCG 780
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
DB 781 TGGCTGTACATCATGTGCCGACATTTGCCAATGCTAGCCCTTGAACACAGCCATC 840
QY 281 AsnPhePheLeuTyrCysPheIleSerValArgPheArgThrMetAlaAlaThrLeu 300
DB 841 AACTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATGCGCCGACGCTC 900
QY 301 LysAlaPhePheLeuCysGlnLeuGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 901 AAGCTTCTCTCAAGTCCAGAGAACCTGTACAGTTCTACACCATATACCTTTTCC 960
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleIleYsMetLeuValTyr 340
DB 961 ATACAGAGTAGCCCTCGATCTCGCGGAAATCTCACTGATCAAGATGCTGGTGC 1020
QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
DB 1021 CAGTATGACAAAATGAAACCTATMAAAGTATCCCG 1059

RESULT 6
BD103852 1062 bp DNA linear PAT 27-ANG-2002
LOCUS BD103852
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD103852
VERSION BD103852.1 GI:22649426
KEYWORDS MO 0194582-A/4.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Terao, Y., Matsui, H. and Shintani, Y.
AUTHORS Novel G protein-coupled receptor and its DNA
TITLE Patent: WO 0194582-A 4 13-DEC-2001;
JOURNAL TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

COMMENT OS Homo sapiens (human)
PN MO 0194582-A/4
PD 13-DEC-2001
PR 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP 00P 170446, 23-JUN-2000 JP 00P 194926 PI
YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00, A61P29/00,
PC A61P9/00, A61P35/00, A61P3/00, A61P37/00, A61P1/00, G01N33/566, PC G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
FT Location/Qualifiers
FT /organism='Homo sapiens (human)'.
/db_xref='taxon:9606'

ORIGIN

Db 152 TGGCGCTGGGGCTTGGGTTTCGTCGCCGTGCTACTACAGCCCTTTCGTCGCTCGGT 211
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
 Db 212 TTACCGACAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAAAG 271
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
 Db 272 TCTCTCTACACATATCTCTGGACCTGCTGCTGGCCAGCATCTTGCTCTTTTCATA 331
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 Db 332 GTGTTGTGGACCTTCCGTTTGGAAATTTCACTTTGAACATGAGATGCCCTCAGGTCCC 391
 QY 101 AspPheIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 Db 392 GACAAATCATATAGAGTGGAAATCTCATCTCATCCACACCTTCATATGATTAAGTCTGTA 451
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuValTyrHisThrValSer 140
 Db 452 CGGTAAACATTAACAGATATATGCTGCTGCTGCCACCCGCTCAAGTACACACAGGCTCA 511
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 Db 512 TACCCAGCCCGCACCCGGAAAGTCATTGTATAGTGTATACATCACTGCTTCTGACCAAGC 571
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 Db 572 ATCCCTATTAACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 Db 632 CACGCTCTCATCTGAGATCCACCTCTTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLysValGly 220
 Db 692 ATCTTAACATCATATGTTGTATCAAGCTCAGAGAGAGAGCAATTTCTCTCCGTGCTG 751
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrLeu 240
 Db 752 TACTCCAGCGGGAGAGACACCGGCATCTGTTTCAACATTAACATCTTGTGCGACACTT 811
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 Db 812 TGGGCCCCCGGCAATCATATGATTTCTTACACCTCTATGAGGCGCCCATCCAGAACCGC 871
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
 Db 872 TGGCTGTACATCATATGCTCCGACATGCTGCAACATGCTGACCTTCTGAAACACAGCCATC 931
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
 Db 932 AACTTCTTCTTACTGCTTCAATCAGCAAGCGGTTCCGACATGACAGCGCCAGCGCTC 991
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 Db 992 AAGGCTTTCTTCAAGTGCAGAAAGCAACCTGTACAGTTCATCAACATCAATCACTTTTCC 1051
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 Db 1052 ATTAACAAGATAGCCCTGGATCTGCGCGGCAAACTCACAATGATCAAGATGCTGCTGAC 1111
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 Db 1112 CAGTATGACAAATAATGAAAACTATTAAGTATCCCG 1150
 RESULT 8
 AX254977 1343 bp DNA linear PAT 10-OCT-2001
 LOCUS AX254977
 DEFINITION Sequence 21 from Patent WO0170978.
 ACCESSION AX254977
 VERSION AX254977.1 GI:16074504
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 AUTHORS Taupier, R. J., Majumder, K., Spaderna, S. K., Smitheon, G., Mezes, P. S.
 and Vernet, C. A.
 TITLE Polypeptides and nucleic acids encoding same
 JOURNAL Patent: WO 0170978-A 21 27-SEP-2001;
 Curagen Corporation (US)
 FEATURES
 source
 1..1343
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,13e-148 Length: 1343
 Score: 1868.00 Matches: 352
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 99.79% Indels: 0
 DB: 6 Gaps: 0
 US-09-995-225B-16 (1-353) x AX254977 (1-1343)
 QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
 Db 2 ATGAGACACACCGACCGCCACCTCGCAGCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61
 QY 21 SerAlaCysGlyLeuGlyLysValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
 Db 62 TGGCCTGGGGCTTGGGTTTCGTCGCCGTGCTTCTTCAACGCTTGTGCTGCTGCTGCTGCT 121
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
 Db 122 TTACCGACAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGAACAGAAAG 181
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
 Db 182 TCTCTCTACATATCTTGGACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 Db 242 GTGTTGTGACCTTCTGTTGAGAGATTCATCTTGAATCAGATGCTGCTGCTGCTGCTGCT 301
 QY 101 AspPheIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 Db 302 GACAAATCATATAGAGTGGAAATCTCATCTCATCCACACCTCCATATGATTAAGTCTGTA 361
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuValTyrHisThrValSer 140
 Db 362 CGGTAAACATTAACAGATATATCACTGCTGCCACCGCTCAGATACACACCGGCTCTCA 421
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 Db 422 TACCCAGCCCGCACCCGGAAAGTCATTGTATAGTTCATCAACCTGCTCTGACCAAGC 481
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 Db 482 ATCCCTATTAACAGTGGGCCCAACATCTGATGTAAGACTCAATCAACACCTCTGTGAT 541
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 Db 542 CACGCTCTCATCTGATCCATGCTTCAACGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLysValGly 220
 Db 602 ATCTTGAACTCATATATGTTGTAACAGCTCAGAGAGAGAGCAATTTTCTGCTCGGTGCT 661
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrLeu 240

Db 662 TACTCAGGGGAGACCAACCGCATCTTTGTCACATTACCTCCATCTTGGCAGACTT 721
Qy 241 TTPALAPROAAGIIEIEMETILEUZYRHISLEUZYRGIALAPROLLEGIANAPARG 260
Db 722 TGGGGCCCCCGCATCATCATGATTTCTTACACCTCATGGGGCCCATCCAGAACCGC 781
Qy 261 TTPLEUVALHISILEMETSERASPILEALASMETLEUALALEUASANTHRAIAILE 280
Db 782 TGGCTGGTACATCATCATGTCGACATGTCGCAACATGCTAGCCCTTCTGAAACACAGCCATC 841
Qy 281 ASNPHEPELEUZYCYSPHEIIESERLYSARGPHEARGTHMETALALATHLEU 300
Db 842 AACTCTCTCTCTCATCTCTCATCAGCAGCGGTTCGCAACATGAGCCGCGCAGCTC 901
Qy 301 LYSALAPHEPELYSCYSGINLYSGINPROVALGINPHERYTHRASNHSANPHESES 320
Db 902 AAGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCATCATCAATCACTTTCC 961
Qy 321 ILETHSERSEPROTPRISESERPROALASNSERHISCYAILLEYSMETLEUVALTYR 340
Db 962 ATACAGAGTACCCCTGGATCTGCGCGCAAACTCACACTGCATCAAGATGCTGTGTAC 1021
Qy 341 GINTYRASPYSANGLYLYSPROLLEYVALSERPRO 353
Db 1022 CAGTATGACAAAATGGAACCTATMAAGTATCCCG 1060

RESULT 9
AX254975 1119 bp DNA linear PAT 10-OCT-2001
LOCUS Sequence 19 from Patent WO0170978.
DEFINITION AX254975
ACCESSION AX254975.1 GI:16074503
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Taupier, R.J., Majumder, K., Spaderna, S.K., Smithson, G., Mezes, P.S.
and Vernet, C.A.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0170978-A 19 27-SEP-2001;
Curagen Corporation (US)
FEATURES
source Location/Qualifiers
1..1119
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 7.79e-148 Length: 1119
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.20% Indels: 0
DB: 6 Gaps: 0

US-09-995-225B-16 (1-353) x AX254975 (1-1119)

Qy 1 MetGluHietHrhieAlHietLeuAlaIaAnseSerLeuSerTrpTrpSerProGly 20
Db 1 ATGGAGCACACGACGCCCACTCCGACGCAACGCTGCTGTGTGTCCTCCCGCGC 60
Qy 21 SerAlaCySgLyLeuGlyPheValProValValTYRYSerLeuLeuCYsLeuGly 40
Db 61 TCGGCTTCGGCTGGCTTCTGTCGTCGCTGCTCACTACACCTTCTGCTGCTGCT 120
Qy 41 LeuProAlaAenIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db 121 TTACCAAGAAATATCTTACAGATATCATCTCTCCACGCTGTCGCAAGACAGAG 180
Qy 61 SerSerTYRAspTYRLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80

Db 181 TCTCTTACACATATCTTGGCAGCTCGCTGTCGCGACATTTGTCTCTTTTCATA 240
Qy 81 ValPheValAspPheLeuGluAspPheIleLeuAenMetGlnMetProGlnValPro 100
Db 241 GTGTGTGTGAGCTTCTGTGTGAGATTTCACTTGAACATGACATGACGCTCCCG 300
Qy 101 AspLYSILEIIEGLVALLEUGLUPHESESERILEHIEPTHSERILETRPILETHRAI 120
Db 301 GACAGATCATGAAAGTGTGAAATTTCAATTCATCATCAACACCTCCATATGATTACTGTA 360
Qy 121 PROLEUTHRIEASPARYTRILEALAYALCYSHISPROLEUZYRHIETHRYALSER 140
Db 361 CCGTTAACCATATGACAGGATATCGCTGTGCGACCGGCTCAAGTACACACGGTCTCA 420
Qy 141 TYRPROALARGYTHRARGLYVALILEVALSERVALTYRILETHRCYSPHELEUTHSER 160
Db 421 TACCAAGCCGACCCGGAAGTCAATGTATGATTTACATCACTGCTTCTTCAACGAC 480
Qy 161 ILEPROTYRITYRTPTRPROASNILETRPHRGUASPTYRILESERTHSERVALHIS 180
Db 481 ATCCCTATTTACTGGTGGCCCAACATCTGACCTGAAGCTACATCAGACCTCTGTGCAT 540
Qy 181 HISVALLEUILETRPILIEHISCYSPHETHRVALTYRLEUVALPROCYSSERILEPHE 200
Db 541 CAGCTCTCATCTGATCTGATCTGCTTCAACCTCTTACCTGTGCGCTGCTCATCTTCTTC 600
Qy 201 ILELEUANSERILEIIEVALTYRLYSLEUARGHLYSERASNPHEARGLEUARGLY 220
Db 601 ATCTTGAACCTGAACTATGTGTACAGCTCAGAGAGAAAGCAATTTTCCTCCGAGGC 660
Qy 221 TYRSETRHGRGLYSTRHTRHALILEUUPHETHRILETHSERILEPHEALATHLEU 240
Db 661 TACTCCAGGGGAGACACCGCATCTGTTCACCATCACTTCATCTTGGCAGACTT 720
Qy 241 TTPALAPROAAGIIEIEMETILEUZYRHISLEUZYRGIALAPROLLEGIANAPARG 260
Db 721 TGGGGCCCCCGCATCATCATGATTTCTTACACCTCATGGGGCCCATCCAGAACCGC 780
Qy 261 TTPLEUVALHISILEMETSERASPILEALASMETLEUALALEUASANTHRAIAILE 280
Db 781 TGGCTGGTACATCATCATGTCGACATGTCGCAACATGCTTACCCCTTCTGAAACAGCCATC 840
Qy 281 ASNPHEPELEUZYCYSPHEIIESERLYSARGPHEARGTHMETALALATHLEU 300
Db 841 AACTCTCTCTCTCATCTCTCATCAGCAGCGGTTCGCAACATGTCGCGCAGCGCAGCTC 900
Qy 301 LYSALAPHEPELYSCYSGINLYSGINPROVALGINPHERYTHRASNHSANPHESES 320
Db 901 AAGGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCAATCACTTTCC 960
Qy 321 ILETHSERSEPROTPRISESERPROALASNSERHISCYAILLEYSMETLEUVALTYR 340
Db 961 ATACAGAGTACCCCTGGATCTGCGCGCAAACTCACACTGCATCAAGATGCTGTGTAC 1020
Qy 341 GINTYRASPYSANGLYLYSPROLLEYVALSERPRO 350
Db 1021 CAGTATGACAAAATGGAACCTATMAAGTATTAADA 1050

RESULT 10
AX375235 1130 bp DNA linear PAT 01-MAR-2002
LOCUS Sequence 25 from Patent WO0210387.
DEFINITION AX375235
ACCESSION AX375235.1 GI:19169953
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
Bilic, V.S., Ramkumar, J., Baughn, M.R., Kallick, D.A., Walla, N.K.,

Hafalla, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L.,
Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.
G-protein coupled receptors
Patent: WO 0210387-A 25 07-FEB-2002;
Incyte Genomics, Inc. (US)

FEATURES

source 1. 1130
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474767CB1"

ORIGIN

Alignment Scores:

Pred. No.: 7.88e-148 Length: 1130
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.20% Indels: 0
DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x AX375235 (1-1130)

QY 1 MetGluHsrThrsAlaHisLeuValAlaAsnSerSerLeuSerTrpTrpSerProGly 20
DB 12 ATGAGACACACGACGCGCCACCTCGACGCCACAGCTCGCTTGGTGGTCCCCGGC 71
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB 72 TCGGCTCGCGGCTTGGTTCGTCCTCGGTCTACTACAGCCCTTGGTGGTCTCGT 131
QY 41 LeuProAlaAsnLeuLeuThrValIleIleLeuSerGlnLeuValAlaAlaGlyGlnLys 60
DB 132 TTACGACGAAATCTTGACAGTATCATCTCCACGCTGGTGGCAGAGACAGAG 191
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
DB 192 TCCTCCACACATCTCTTGGCAGCTGCTGCTGCCACATCTTGGCTCTTTTATTA 251
QY 81 ValPheValAspPheLeuLeuGlyAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 252 GTGTTGTGACCTTCTGTTGAGAGATTTTCATCTTGAACATGACGATGCCAGGTCCC 311
QY 101 AspArgIleIleGluValLeuGlyPheSerSerIleIleThrsSerIleTrpIleThrVal 120
DB 312 GACAAAGATCATAGAGAGCTGGAATTCATCTCCACACCTCCATATGATTAATCTGA 371
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 372 CCGTTACCATGACAGATATATGCTGCTGCCACCGCTCAAGTACACACGCTCTCA 431
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 432 TACCCAGCCGCGACCCGGAAGATCATGTAAGTTTACATCACCTGCTCTGACACAGC 491
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
DB 492 ATCCCTCATTAATCGTGTGCCCAACATCTGACGAAACATACATACACACCTCTGTCAT 551
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
DB 552 CAGCTCTCATCTGATCCACTGCTTCAACGCTTACCTGAGCCCTGCTCATCTTCTTC 611
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
DB 612 ATCTTGAATCATCATATTGTGTACAGCTCAGAGAGAGCAATTTTCGTCCTCGTGGC 671
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 672 TACTCCAGGGGAGACACCGCATCTTGTTCACATTAACCTTCATCTTGGCACACTT 731
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260

DB 732 TGGGCCCCCGCATCATCATGATTTCTTACCACTCTATGAGGGGCCCATCCAGAACCGC 791
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
DB 792 TGGCTGTATACATCATATGTCGACATTTGCCAATCATGATGCCCTTGTGAACAGACCATC 851
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetIleAlaIleThrLeu 300
DB 852 AACCTTCTCTTCACTGCTTCATACAGCGGATTCGACCAATGGACGCGCACGCTC 911
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 912 AAGCTTCTTCTCAAGTCCAGAGCAACCTGTACAGTCTACACCATCATTAATTCTTC 971
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
DB 972 ATAAACAAGTACCCCTGATCTCGCGGCAATCTACACATGACATCAAGATGCTGTAC 1031
QY 341 GlnTyrAspLysArgLysArgProIleLys 350
DB 1032 CAGTATGACAAATAAGAAAACCTATPAAA 1061

RESULT 11

AX664703 1062 bp DNA linear PAT 22-MAR-2003
LOCUS
DEFINITION Sequence 12 from Patent WO02074960.
ACCESSION AX664703
VERSION AX664703.1 GI:29164463
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS

Leiby, K.R., Kapeller-Libermann, R. and Glucksmann, M.

38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof

TITLE

Patent: WO 02074960-A 12 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)

JOURNAL

Location/Qualifiers

FEATURES

source 1. 1062
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 9.22e-147 Length: 1062
Score: 1844.00 Matches: 347
Percent Similarity: 98.87% Conservative: 2
Best Local Similarity: 98.30% Mismatches: 4
Query Match: 98.50% Indels: 0
DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x AX664703 (1-1062)

QY 1 MetGluHsrThrsAlaHisLeuValAlaAsnSerSerLeuSerTrpTrpSerProGly 20
DB 1 ATGAGACACACGACGCGCCACCTCGACGCCACAGCTCGCTTGGTGGTCCCCGGC 60
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB 61 TCGGCTCGCGGCTTGGTTCGTCCTCGGTCTACTACAGCCCTTGGTGGTCTCGT 120
QY 41 LeuProAlaAsnLeuLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
DB 121 TTACGACGAAATCTTGACAGTATCATCTTCCACGCTGCTGCAAGACAGAGAG 180
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
DB 181 TCCTCTTACAAATATCTTGGCAGCTGCTGCTGCCACATCTTGGTCTCTTTTATTA 240
QY 81 ValPheValAspPheLeuLeuGlyAspPheIleLeuAsnMetGlnMetProGlnValPro 100

```

Db      241 GGTGTTGGAGCTTCTCTGGAGAGATTTCATCTTGAACATGCGAGATGCCAGGTCCCC
Qy      101 AsplysillellegluValleugluPheserSerillesthrSerillethrpillethrpVal 120
Db      301 GACAAAGATCAATGAAAGTCTGGAAATTCATCCATCCACACTCCATATGATGATTAAGTGA 360
Qy      121 ProleuthrilleapargTyrillealavalCysHisProleuThyThrsValSer 140
Db      361 CGGTAAACATTGAGAGGTATATGCGTGTGTCACACCGCTCAAGATCAACACCGTCTCA 420
Qy      141 TyrProalaargThrargValleValserValTyrillethrpCysPheleuthrSer 160
Db      421 TACCAGCCCGCACCCGGAAAGCATGTTGATGATTATACACTGCTCTCCAGACGC 480
Qy      161 IleProtyrTyrTrpTrpProAsnilleTrpThrGluAspTyrilleSerthrsValHis 180
Db      481 ATCCCTATTATCTGTGGCCCAACATCTGAGCTGAGACTAGACATACATACACCTGTGCAT 540
Qy      181 HisValleuilletPilleHisCysPheThrpValTyrleuValProCysSerillePhePhe 200
Db      541 CAGCTCCTCATCTGGGTCCACTGCTTCAACGCTTACACAGGTCCTGCTCCATCTCTTC 600
Qy      201 IleleuAsnSerilleleValTyrLysleuArgArgLysSerAsnPheArgleuArgGly 220
Db      601 ATCTTGAACATCAATCGTTGTGTACAAAGCTCAGAGAGAAAGCAATTTTGTCTGTGGC 660
Qy      221 TyrSerThrGlyLysThrThralalleleuPheThrillethrserillePhealathrleu 240
Db      661 TACTCCACGGGGAAGACACCGCCATCTTGTTCACATTCATCTTCATCTTGGCATACTT 720
Qy      241 TyrAlaProargIleIleMetIleleuTyrHisleuTyrGlyAlaProIleGlnAsnArg 260
Db      721 TGGGCCCCCGGCATCATCATGATTTCTTACCATCTGATGGGGCGCCATCCGAACCGC 780
Qy      261 TrpLeuValHisIleMetSerAspIleAlaAsnMetleuAlaIleuAsnthralalle 280
Db      781 TGGCTGTATACATCATGTCGACATTCGCCAACATCTAGCCCTCTGAAACACAGCCATC 840
Qy      281 AsnPhePheleuTyrCysPheIleSerLysArgPheArgThrMetAlaIleAlaThrleu 300
Db      841 AACTCTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGACCATGACACCGCCACAGCTC 900
Qy      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db      901 AAGGCTTTCTTCAAGGCCAGAGAACCTGTACAGTTCTACACCAATCATTAATTTTCC 960
Qy      321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetleuValTyr 340
Db      961 ATTAACAAGTAAAGCCCTCGATCTCGCCGCAAACTCACACTGCAATCAAGATGCTGTGAC 1020
Qy      341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db      1021 CAGTATGACAAAATGGAATACTATATAAAGTATCCCCG 1059

```

RESULT 12
AX664701
LOCUS AX664701 1526 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 10 from Patent WO02074960.
ACCESSION AX664701
VERSION AX664701.1 GI:29164461
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Leiby, K.R., Kapeller-Libermann, R. and Glucksman, M.
TITLE 38550, 28472, 5495, 65507, 81588 and 14354 methods and compositions
JOURNAL of human proteins and uses thereof
FEATURES Patent: WO 02074960-A 10 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers

```

source 1..1526
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAD80046.1"
        /db_xref="GI:29164462"
        /translation="MEHTHAHLAANSLSWMSPGSACGLGFPVIVYSLQCLGLPAN
        ILTVILISQVLARKQSSYNVLAALAADILVLFPIVDFLBDPILNMQPVPDK
        ILLEVFSSIHSTISWITVPLIDRYAVCHPXYVSPATKRVIVSYITCLTS
        LPYVWPNIMWEDYSTSVHVLVWHPFTVYQVCSIFPTINSIVYKLRKSFRL
        CGYSTGKTALIFTTTSLFALIMAPRITIMILYHYGAPIDRMVIMSDIANNALIL
        NTAINEPLFCFSIKSRFTMAAATLKAFKCKQKQPVQFTNNPFSITSSPWLSPANSHC
        IKMLVYQYDKNCKPKVSP"

Alignment Scores:
Pred. No.: 1..4e-146 Length: 1526
Score: 1844.00 Matches: 347
Percent Similarity: 98.87% Conservative: 4
Best Local Similarity: 98.30% Mismatches: 2
Query Match: 98.50% Indels: 0
DB: 6 Gaps: 0

US-09-995-225B-16 (1-353) x AX664701 (1-1526)
Qy      1 MetGluHisThrHisAlaHisIleuAlaAlaAsnSerSerleuSerTyrTrpSerProGly 20
Db      139 ATGAGACACACGACGACGCCACCTCGACCAAGCTCGTGTGTGGTGTCCCGCCG 198
Qy      21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerleuLeuLeuCysLeuGly 40
Db      199 TCGGCTCGCGCTTGGGTTTGTGTCCTCGGTGTACTACAGCCTCTTGACGTGCTGCT 258
Qy      41 LeuProAlaAsnIleleuThrValIleIleleuSerGlnleuValAlaArgArgGlnLys 60
Db      259 TTACAGCAAAATATCTTGACAGTATATATCTTCCACGCTGGGCGCAAGACAGCAAG 318
Qy      61 SerSerTyrAsnTyrleuLeuAlaAlaAlaAspIleleuValleuPhePheIle 80
Db      319 TCCTCTACACATATCTCTTGAGCACTGCTGCTGCGCAKATCTTGCTCTTTTCATA 378
Qy      81 ValPheValAspPheleuLeuGluAspPheIleleuAsnMetGlnMetProGlnValPro 100
Db      379 GTGTTGTGACCTTCCGTGGAGAGATTTCATCTTGAAACATGCAAGATGCTCAGTCCC 438
Qy      101 AsplysillellegluValleugluPheserSerillesthrSerillethrpillethrpVal 120
Db      439 GACAAAGATCAATGAAAGTCTGGAAATTCATCCATCCACACTCCATATGATGATTAAGTGA 498
Qy      121 ProleuthrilleapargTyrillealavalCysHisProleuThyThrsValSer 140
Db      499 CGGTAAACATTGAGAGGTATATGCGTGTGTCACACCGCTCAAGTACACACAGGTCTCA 558
Qy      141 TyrProalaargThrargValleValserValTyrillethrpCysPheleuthrSer 160
Db      559 TACCAGCCCGCACCCGGAAAGCATGTTGATGATTATACACTGCTCTCCATCTCTTC 618
Qy      161 IleProtyrTyrTrpTrpProAsnilleTrpThrGluAspTyrilleSerthrsValHis 180
Db      619 ATCCCTATTATCTGTGGCCCAACATCTGAGCTGAGACTAGACATACACCTGTGCTCAT 678
Qy      181 HisValleuilletPilleHisCysPheThrpValTyrleuValProCysSerillePhePhe 200
Db      679 CAGCTCCTCATCTGGGTCCACTGCTTCAACGCTTACACAGTCCCTGCTCCATCTCTTC 738
Qy      201 IleleuAsnSerilleleValTyrLysleuArgArgLysSerAsnPheArgleuArgGly 220
Db      739 ATCTTGAACATCAATCGTTGTGTACAAAGCTCAGAGAGAAAGCAATTTTGTCTGTGGC 798
Qy      221 TyrSerThrGlyLysThrThralalleleuPheThrillethrserillePhealathrleu 240

```

Db 799 TACTCCAGGAGGAGACCGCATCTTGTACCACTTACCTTCATCTTGTCCATCTT 858
Qy 241 TTPALAFroArlgilelleMeclleuYrhisleuYrGlyalProleginaAarg 260
Db 859 TGGGCCCCCGCATCATCATGATCTTTTACCACTTATGGGGGCCCATCCAGAACCGC 918
Qy 261 TTPLeuValHisileMeSerAapIleAlaAmMeclleuAlaleuLeuAenThraIle 280
Db 919 TGGCTGTATACATCATCATGCTCGACATTGCCAACATGTCTACCCCTTCGAACACAGCATC 978
Qy 281 AsnPhelheleuYrCySpheileSerlyeArpPheArghThrMetAlaAlaathleu 300
Db 979 AACT 1038
Qy 301 LysAlaPhePheleuYrCySpheileSerlyeArpPheArghThrMetAlaAlaathleu 320
Db 1039 AAGGCTTTCTTCAAGTCCAGAACCACTGTACAGTTTACACCATCATTAATTCTTTC 1098
Qy 321 IleThrSerSerProTPIleSerProAlaAmSerHisCySyllelyeMeclleuVal 340
Db 1099 ATACACAGTACCCCGCTGATCTCGCCGCAAACTCACATGACATCAAGATGTGTGTAC 1158
Qy 341 GluTyraAspLysAsnGlyLysProIleValSerPro 353
Db 1159 CAGTATGACAAATATGAAATCTTATTAAGTATCCCG 1197

RESULT 13
AX463235 1062 bp DNA linear PAT 15-JUL-2002
LOCUS
DEFINITION
Sequence 9 from Patent WO0246358.
ACCESSION
AX463235.1 GI:21886203
VERSION
KEYWORDS

SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Okigami, H.
Regulation of human chemokine-like receptor
Patent: WO 0246358-A 9 20-JUN-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1,47e-138 Length: 1062
Score: 1747.00 Matches: 335
Percent Similarity: 90.81% Conservative: 1
Best Local Similarity: 90.54% Mismatches: 0
Query Match: 93.32% Indels: 34
DB: 6 Gaps: 2

US-09-995-225B-16 (1-353) x AX463235 (1-1062)

Qy 1 MetGluHleThrlaHisleuAlaAlaAmSerleuSerTTPTPSerProGly 20
Db 1 ATGAGACACACGACCGCCACCTCGACGCAACAGCTGTGCTTGGTGTCCCGGC 60
Qy 21 SerAlaCyegIleuGlyPheValProValValTyTySerleuLeuCySleuGly 40
Db 61 TCGGCGCGCGCTTGGTTCGTGCGCGCTGTCTACACAGCTCTTGTCTGTCTCGT 120
Qy 41 LeuProAlaAmIleleuThraValIleleuSerGlnleuValAlaArgArgGlnLys 60
Db 121 TTACCGACAAATATCTTGAACAGTATCATCTCTCCACGCTGTGTGCAAGACAGAA 180
Qy 61 SerSerTyraenTyrlleuLeuAlaAlaAlaAspIleleuValleuPhePheIle 80

Db 181 TCTCTTACACATCATCTTGTGACATCGCTGCGCCGACATCTTGTCTCTTTTCATA 240
Qy 81 ValPheValAspPheleuGluAapPheIleleuAmMeclleuMetProGlnValPro 100
Db 241 GTGTGTGTGACTTCTGTGTGAAGATTTTCACTTTGAACATGACATGCTCAGTCCCC 300
Qy 101 AspLysIleIleGluValleuGluPheSerSerIleHisThrSerIleTPIleThraVal 120
Db 301 GACAAATACATTAAGATGTGGAATTTCTCATCATCACACCTCCATATGATTAAGTGA 360
Qy 121 ProleuThrIleAspArgTyrlleAlaValCySHIaProleuYrhisThraValSer 140
Db 361 CGCTTAACCTTACACAGATATATCGCTGTCTGACACCGCTCAAGATACACAGCTCTCA 420
Qy 141 TyrProAlaArgThraArgValIleValSerValTyrlleThrCySpheleuThraSer 160
Db 421 TACCCAGCCCGCAC-----ValHisHisValleu 435
Qy 161 IleProTyTyTyTTPTPProAlaenIleTTPThrgluAapTyrlleSerThraSer 178
Db 436 -----TATTAATGATGAGCCCAACATCTGATGAAAGACTATACACACTCTGCCCGG 489
Qy 179 -----ValHisHisValleu 183
Db 490 AAGATCATTTGAAGTGTATACATCACCTGCTCTGACACAGCATCTGACATCGTCTTC 549
Qy 184 IleTPIleHisCySpheThraValTyrlleuValProCySerIlePhePheIleleuAsn 203
Db 550 ATCTGATCCATGCTTCAACCGCTTACCTGTGTCCTGCTGCTCATCTTCTTCACTTGAAC 609
Qy 204 SerIleIleValTyrlleleuArgArgLysSerAsnPheArgleuArgGlyTyrlleThra 223
Db 610 TCAATCATTTGTACAGCTCAGAGAAAGAGCAATTTTCCTCGTGTGTACTCCAG 669
Qy 224 GlyLeuThraValIleleuPheThrlleThraSerIlePheAlaThraIleuThraPro 243
Db 670 GGGAAAGACACCGCCATCTTGTTCACCATTACTTCATCTTGGACACTTGGGCCCC 729
Qy 244 ArgIleIleMeclleuYrhisleuYrGlyalProIleGlnAmArgTPIleuVal 263
Db 730 CGCATCATCATGATTTCTTACCACTCTATGGGGGCCCATCCAGAACCGCTGTGTGA 789
Qy 264 HisIleMeSerAspIleAlaAmMeclleuAlaleuLeuAenThraIleleuPhePhe 283
Db 790 CACATCATGTCCGACATTCGCAACATGATGATGCTTGTGAACACACCTCATCTTCTTC 849
Qy 284 LeuTyCySpheIleSerlyeArpPheArghThrMetAlaAlaathleuValAlaPhe 303
Db 850 CTCTACTGCTTCATACAGAGCGGTTCCGACCATGACGCGCCACGCTCAAGGCTTTC 909
Qy 304 PheLysCyegIleuGluPheValGlnPheTyThraAmHisAsnPheSerIleThraSer 323
Db 910 TTCAAGTGCAGAGAACCTGTATGACCTTACACCAATCATTAATTTCATTAACAAAT 969
Qy 324 SerProTPIleSerProAlaAmSerHisCySyllelyeMeclleuValTyrlleThra 343
Db 970 AGCCCTGATCTCGCCGCAAACTCACACTGATCAAGATGTGTGTACAGTATGAC 1029
Qy 344 LysAsnGlyLysProIleValSerPro 353
Db 1030 AAAAATGAAATCTTATTAAGTATCCCG 1059

RESULT 14
CO737667 957 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION
Sequence 23601 from Patent WO02066579.
ACCESSION
CO737667
VERSION
CO737667.1 GI:42336565
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
PE Corporation (NY) (US)
06-SEP-2002;
Patent: WO 02068579-A 23601
kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof

FEATURES
Source
Location/Qualifiers
1. .957
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 8,18e-130 Length: 957
Score: 1643.00 Matches: 312
Percent Similarity: 99.36% Conservative: 0
Best Local Similarity: 99.36% Mismatches: 2
Query Match: 87.77% Indels: 0
DB: 6 Gaps: 0

US-09-995-225b-16 (1-353) x CQ737667 (1-957)

QY 40 G1YLeuPProAlaAenIleuThrValIleIleuSerGlnleuValAlaArgGln 59
Db 13 GGTTCGACGAAATATCTTGACGATGATCATCTCTCCGAGCGGGAAGACAG 72
QY 60 LysSerSerTyraNtyrIleuLeuAlaLeuAlaAlaAspIleuValIleuPhe 79
Db 73 AAGTCCCTCAACAATCTCTTGACGATGATCATCTCTCCGAGCGGGAAGACAG 132
QY 80 IleValPheValAspPheLeuLeuGluAspPheIleuAsnMetGlnMetProGlnVal 99
Db 133 ATAGTGTTCGACCTTCCTGTTGAAAGATTATCTTGAAACATGAGATGCTCAGGTC 192
QY 100 ProAspLysIleIleGluValIleuGluPheSerSerIleIleTherSerIleTrrIleThr 119
Db 193 CCGGACAGATCATAGAGTGTGGAATCTCATCATCATCATCATCATCATCATCATCAT 252
QY 120 ValIleuThrIleAspArgTrrIleAlaValCysHisProLeuLysTrrIleThrVal 139
Db 253 GATCCGTTAACCTTACAGATATATGCTGCTGACCCCTCAAGTACACAGCGGTC 312
QY 140 SerTyrProAlaArgThrArgLysValIleValSerValTrrIleTherCysPheLeuThr 159
Db 313 TCATACCCAGCCCGGACCCGGAAGTATGATGTTTATCATCATCATCATCATCATCAT 372
QY 160 SerIleProTyrTrrTrrProAsnIleTrrPheGluAspTrrIleSerTherSerVal 179
Db 373 AGCATCCCTATATACGTCGTCGCAACATCTGAGCTGAAGACTATCATCATCATCATCAT 432
QY 180 HisHisValIleuIleTrrIleHisCysPheThrValTrrIleValProCysSerIlePhe 199
Db 433 CATCATCTCTCATCTGATTCACGCTGCTACCGTCACTGTCCTGCTCCCATCTTC 492
QY 200 PheIleLeuAsnSerIleIleValTrrLysLeuAlaGArgLysSerAsnPheArgLeuArg 219
Db 493 TTCATCTTGAATCATCATCTGTTGTAACAAGCTCAAGAGGAAGAGAAATTTTCGTCCT 552
QY 220 G1YTrSerThrGlyLysThrThrAlaIleuPheThrIleTherSerIlePheAlaThr 239
Db 553 GGTACTCTCAGCGGGAAGACACCGCATCTTGTTCACCATTAACCTCATCTTGGCCACA 612
QY 240 LeuTrrAlaProArgIleIleMetIleuLysIleuValIleuValProIleGlnAsn 259
Db 613 CTTTGGGCCCCCGCATCATCATCTTTCACCTCTTATGGGGCCCATCCAGAAC 672
QY 260 ArgTrrPheValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAla 279
Db 673 CGTGGCTGTACATCATGTCGACATGTCGAACATGTCAGCCCTTCTGAACACAGCC 732

QY 280 IleAsnPhePheLeuTrrCysPheIleSerIleLysArgPheArgThrMetAlaAlaThr 299
Db 733 ATCAATCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 792
QY 300 LeuLysAlaPhePheLysCysGlnLysGlnProValGlnPheTrrThrasnHisAsnPhe 319
Db 793 CTCAGGCTTCTTCTTCAAGTCCAGAAAGCACTGATGTTTATCAACCATCATCATCTT 852
QY 320 SerIleTherSerProTrrIleSerProAlaAsnSerHisCysIleLysMetLeuVal 339
Db 853 TCATACAGATGAGCCCGGATCTGCGCGGAACTCATCATCATCATCATCATCATCATCAT 912
QY 340 TrrGlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db 913 TACAGTATGACAAATAAGAAATGAAATTAAGATATCCCG 954

RESULT 15
CQ739127 1788 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION
Sequence 25061 from Patent WO02068579.
ACCESSION
CQ739127
VERSION
CQ739127.1 GI:42341363
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
PE Corporation (NY) (US)
06-SEP-2002;
Patent: WO 02068579-A 25061
kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof

FEATURES
Source
Location/Qualifiers
1. .1788
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,24e-129 Length: 1788
Score: 1641.50 Matches: 320
Percent Similarity: 91.32% Conservative: 6
Best Local Similarity: 89.64% Mismatches: 10
Query Match: 87.69% Indels: 21
DB: 6 Gaps: 2

US-09-995-225b-16 (1-353) x CQ739127 (1-1788)

QY 9 AlaAlaAsnSerSerLeuSerTrrTrrPheProGlySerAlaCysGlyLeuGlyPheVal 28
Db 742 TCAGCTCACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 777
QY 29 ProValValTrrTrrSerLeuLeuCysLeuGly----- 40
Db 778 ---ATCTCTCAACGTTCTCGATGTTGCTGCGGGAATCTACCTGTACAAAGAAATG 834
QY 41 -----LeuProAlaAenIleuThrValIleIleuSerGlnleuValAla 56
Db 835 GGAACCTCAGTGTCTGCGGGAATATCTTGACGATGATCATCTCTCCAGCTGTGAGCA 894
QY 57 ArgArgGlnLysSerSerTrrTrrIleuLeuAlaLeuAlaAlaAspIleuVal 76
Db 895 AGAAGACAGAAAGCTCCCTTCAACATATCTTGGGCACTGCTGCGGCAACATCTTGGTC 954
QY 77 LeuPhePheIleValPheValAspPheLeuGluAspPheIleuAsnMetGlnMet 96
Db 955 CTTTTCATATGTTTGTGATCTTCTGTTGAAAGATTATCTTGAAACATGAGATG 1014
QY 97 ProGlnValProAspLysIleIleGluValIleuGluPheSerSerIleIleTherSerIle 116

Db 1015 CCTCAGTCCCGACAGATCATAGAGTGTGGAATTCATCCATCCACACTCCATA 1074
Qy 117 TTPLEThValProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyr 136
Db 1075 TGGATTACTGTAACCGCTTAACCATTTGACAGATATATGCTGTGCGACCCGCTCAAGTAC 1134
Qy 137 HistHValSerTyrProIleArgThrArgLysValIleValSerValTyrIleThrCys 156
Db 1135 CACAGGCTCATACCCAGCCGACCCGAGAAAGTCATTGTAAGTGTATTAATCATCACTGC 1194
Qy 157 PheLeuThrSerIleProTyrTyrTrpIleProAsnIleTrpThrGluAspTyrIleSer 176
Db 1195 TTCCTGACAGCATCCCTATTACTGTGGCCCAACATCTGAGCTGAGATCATCACTGAC 1254
Qy 177 ThrSerValHisHisValIleuIleTrpIleHisCysPheThrValTyrLeuValProCys 196
Db 1255 ACCCTGTGTCATCACGCTCTCATCTGTGATCCACTGCTCACCGTCACTGAGGCGCTGC 1314
Qy 197 SerIlePhePheIleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPhe 216
Db 1315 TCCATCTTCTTCACTTGAATCATATGTTGATACAGCTCAGAGAGAGCAATTTT 1374
Qy 217 ArgLeuArgGlyTyrSerThrGlySerThrAlaIleLeuPheThrIleThrSerIle 236
Db 1375 CGTCTCGGTGCTACTCCACGGGAGAACACCGCATCTTGTTCACCATTACTTCATC 1434
Qy 237 PheAlaThrLeuTrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaPro 256
Db 1435 TTGGCACACTTTGGGCCCCCGCATCATGATGATTTTACCACTTATGGGGCGCCC 1494
Qy 257 IleGlnAsnArgTrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeu 276
Db 1495 ATCCAGAAACCGTGGCTGTACATCATGTCGACATTGCCAAATGCTAGCCCTTCG 1554
Qy 277 AsnThrAlaIleAsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAla 296
Db 1555 AACACAGCATCACTTCTTCTCTACTGCTTCAACGCAAGGGTTCCGACCATGGCA 1614
Qy 297 AlaAlaThrLeuLysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsn 316
Db 1615 GCCGCCACGCTCAAGGCTTCTTCAAGTCCAGAAACCACTGTACAGTTCTACACCAAT 1674
Qy 317 HisAsnPheSerIlePheSerSerProTyrPheSerProIleAsnSerHisCysIleLys 336
Db 1675 CATACTTTTCAATAACAGTAGCCCTGAGATCTGCCGCAAACTCACACTGCATCAAG 1734
Qy 337 MetLeuValTyrGlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db 1735 ATGCTGGTGTACAGTAGACAAATAATGAAAACTATATAAGTATCCCCG 1785

Search completed: December 14, 2004, 21:34:44
Job time : 3563 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:25:49 ; Search time 411 Seconds
(without alignments)
4508.627 Million cell updates/sec

Title: US-09-995-225B-16
Perfect score: 1872
Sequence: 1 MEHTHAHLANSLSWMSPG.....CIKMLVYQYDKNGKPIKVP 3533

Scoring table:		
BLOSUM62		
Xgapop 10.0	Xgapext 0.5	
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

```
Searched:      4134886 segs, 2624710521 residues
Total number of hits satisfying chosen parameters:  8269772
```

Maximum DB seq length: 20000000000

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

```

Command line parameters:
-MODEL=frame.p2n.model -DEV=xh
-O=/csm2.1/USPTO.spool1/OPM9995225/runcat.14122004.110732.18012/app.query.fasta_1.519
DB=N Geneset.235ep04 -OPM=fastap -SUFFIX=ring -MINMATCH=0 -1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonume2 -TRANS=numan40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09995225.OCN=1.352.OTRUNC=14122004.110732.18012 -KCPU=6 -ICPU=3
-NO MPAP -LARGEBUDERY -NEG_SCORES=0 -WAIT -DSDBLOCK=100 -LONG=
DEV.TIMOUT=120 -WARN.TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : N_Geneseq_23Sep04.*
1:  geneseqn1980s.*
2:  geneseqn1990s.*
3:  geneseqn2000s.*
4:  geneseqn2001as.*
5:  geneseqn2001bs.*
6:  geneseqn2002as.*
7:  geneseqn2002bs.*
8:  geneseqn2003as.*
9:  geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysts of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	1872	100.0	1059	6	ABN84269	Abn84269 Human che
2	1872	100.0	1062	6	ABK15562	ABK15562 CDNA enco
3	1872	100.0	1062	6	ABK15563	CDNA enco
4	1872	100.0	1062	6	ABT04873	Human G p
5	1872	100.0	1062	10	ACC44115	Human AKO
6	1872	100.0	1062	10	ACC17185	Human G p

7	1872	100.0	1702	8	ABV73364	ABV73364 Human TGR
8	1872	100.0	1702	8	ADP70587	ADP70587 Ophiophaga hannah
9	1868	99.8	1343	4	AA515731	AA515731 DNA encod
10	1868	99.8	1343	4	ADU87768	ADU87768 G-couple
11	1868	99.8	1343	12	ADU179324	ADU179324 NOV1l ccc
12	1868	99.8	1343	12	AD056003	AD056003 DNA encoc
13	1868	99.8	2273	12	AD028955	AD028955 Human nov
14	1857	99.2	1119	4	AA515730	AA515730 DNA encod
15	1857	99.2	1119	10	ADU87766	ADU87766 G-couple
16	1857	99.2	1119	12	ADU179322	ADU179322 NOV10a cc
17	1857	99.2	1119	12	AD056001	AD056001 DNA encoc
18	1857	99.2	1130	6	AA0296672	AA0296672 Human G-g
19	1857	99.2	2198	9	ACF05275	ACF05275 Human G-g
20	1844	98.5	1526	8	AA046658	AA046658 Human 7T
21	1844	98.5	1526	8	ACA60889	ACA60889 Human cDN
22	1844	98.5	1526	10	AB557024	AB557024 cDNA encoc
23	1841	98.3	1110	9	ACF05276	ACF05276 Human G-g
24	1841	98.3	2189	10	ADU18120	ADU18120 Human G-g
25	1756	93.8	1038	8	ABV73373	ABV73373 Mouse TGR
26	1756	93.8	1038	12	AD028957	AD028957 Mouse nov
27	1747	93.3	1062	6	ABN84273	ABN84273 Human che
28	1643	87.8	957	12	ADU179347	ADU179347 NOV10b cc
29	1643	87.8	957	12	AD056026	AD056026 DNA encoc
30	1638	87.5	1002	6	AA027501	AA027501 Human G-g
31	1638	87.5	1002	10	AA061658	AA061658 Human GPC
32	1638	87.5	1032	6	ABN84271	ABN84271 Human che
33	1638	87.5	1070	6	ABN84270	ABN84270 Human che
34	1638	87.5	1158	6	ABE242880	ABE242880 Human GPC
35	1638	87.5	1826	6	ABN84272	ABN84272 Human che
36	1631	87.1	1000	6	AA028102	AA028102 Human thyl
37	1607	85.8	946	10	ADU87879	ADU87879 G-couple
38	1471	78.6	864	4	AA542811	AA542811 Human G P
39	1447	77.3	867	6	AA598134	AA598134 Human DNA
40	1447	77.3	894	6	AA598143	AA598143 Human DNA
41	1421	75.4	963	10	ADU12695	ADU12695 Human GPC
42	1317	70.4	795	6	AA598071	AA598071 Human DNA
43	1317	70.4	795	8	ABE42552	ABE42552 Human G P
44	912	48.7	930	4	AA542808	AA542808 Human G P
45	742	39.6	1023	12	AD028937	AD028937 Mouse nov

ALIGNMENTS

	RESULT 1
ABN84269	
ID	ABN84269 standard; cDNA; 1059 BP.
XX	
AC	ABN84269;
XX	
DT	23-SEP-2002 (first entry)
XX	
DE	Human chemokine-like receptor cDNA.
XX	
KM	Chemokine-like receptor; G-protein coupled receptor; receptor; human;
KM	HIV infection; cardiovascular disease; asthma;
KM	chronic obstructive pulmonary disease; candida; antiasthmatic;
KM	vasotrophic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;
KM	antiinflammatory; antiallergic; immunomodulator; gene therapy; gene; ss
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CDS	1..1059
FT	/*tag= a
FT	/partial
FT	/product= "Chemokine-like receptor"
FT	/note= "the CDS does not include a stop codon"
XX	
PN	WO200248358-A2.
XX	
PD	20-JUN-2002.
XX	
PF	12-DEC-2001; 2001WO-EP014571.

XX 14-DEC-2000; 2000US-0255150P.
 PR 02-APR-2001; 2001US-0280110P.
 PR 21-JUN-2001; 2001US-0299474P.
 XX (FARB) BAYER AG.
 XX Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;
 XX MPI, 2002-547858/58.
 DR P-PSDB; ABB79517.
 XX
 PT New isolated polynucleotide encoding a chemokine-like receptor
 PT polypeptide for treating e.g. asthma, myocardial infarction, human
 PT immunodeficiency virus infection, or chronic obstructive pulmonary
 PT disease.
 XX
 PS Claim 1; Fig 1; 114pp; English.

CC The present sequence is that of cDNA encoding a novel human chemokine-
 CC like receptor (see AMN79517) of 353 amino acids and having 7 putative
 CC transmembrane domains, consistent with the structure of a G-protein
 CC coupled receptor. Its closest human homologue is C-C chemokine receptor
 CC 3. The novel receptor is expressed at low levels in most tissues. It is
 CC expressed at a high level in phytohemagglutinin-stimulated CD8+ cells,
 CC but in none of the other immune cells tested. It may act as a receptor of
 CC chemottractant molecules on activated lymphocytes and be involved in
 CC cell trafficking and homing to sites of infection, inflammation or tissue
 CC injury. Regulation of activity of the novel receptor can therefore be
 CC used to treat cardiovascular, immunological and inflammatory diseases,
 CC including asthma and chronic obstructive pulmonary disease (COPD). The
 CC receptor may also be a target for viruses that reside in the nervous
 CC system. Regulating the binding of ligands, e.g. chemottractant molecules
 CC or virus particles, to the receptor can therefore be used to modulate the
 CC immune response to inhibit viral infections, including HIV infection. A
 CC claimed method of reducing activity of the receptor involves contacting a
 CC cell with a reagent (preferably an antibody, antisense oligonucleotide or
 CC ribozyme) to a product (preferably RNA or a polypeptide) encoded by a
 CC polynucleotide encoding the human chemokine-like receptor in vivo or in
 CC vitro. A claimed method of treating a chemokine-like receptor dysfunction
 CC related disease selected from HIV infection, a cardiovascular disorder,
 CC asthma or COPD uses a reagent that modulates a function of the receptor.
 CC
 XX
 SQ Sequence 1059 BP; 240 A; 342 C; 206 G; 271 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.8e-179 Length: 1059
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-995-225B-16 (1-353) x ABN84269 (1-1059)

QY 1 MetGluHisThrHisAlaHisIleuAlaIaAsnSerIleuSerTrpIleProGly 20
 DB 1 ATGGAGCAACAGCCAGCCAGCCAGCCAGCCAGCCAGCTGCTGCTGCTGCCCGC 60
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly 40
 DB 61 TCGGCTCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
 QY 41 LeuProAlaAsnIleLeuThrValIleIleIleuSerGlnLeuValAlaArgGlnLys 60
 DB 121 TTACCAAGAAATATCTTACAGATGATCATCTCCGCCGCTGGGGAAGACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaIaAspIleLeuValIleuPhePheIle 80
 DB 181 TCCCTCTACACTATCTCTTGGCACTCGCTCTCCCAATCTTGGCTCTTTTATA 240
 QY 81 ValPheValaAspPheLeuLeuGlnaAspPheIleLeuAsnMetGlnMetProGlnValPro 100

DB 241 GTGTTGTGGACTTCCTGTGGAAAGATTTCATCTTGAAATGACATGCTCAGGTC 300
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 301 GAAGAATCATAGAAATGTCTGAAATTCATCATCCACACCTCCATATGATTA 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 361 CGGTAAACATTGACAGAGTATATCCCTGTCCACCCGCTCAATGATACACAGGTC 420
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGACCCGGAAGTCAATGTATAGTTTACATCACTGCTTCGACGAC 480
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpHisGluAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTTATCTAGTGGCCCAACATCTGAGCTGAAGACTACATCAGCAGCCTGTG 540
 QY 181 HisValIleuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CACGTCCTCATCTGGATCCATCGTTCAACGCTTACCTGTGCTGCTCATTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrIleLeuArgArgLysSerAsnPheArgLysGly 220
 DB 601 ATCTTGAACCTCAATCATGTGTACAGCTCAGAGGAAGAACAAATTTTCCTCGTGC 660
 QY 221 TyrSerThrGlyLysThrThrAlaIleuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCAGCGGGAAGACACAGCCATCTGTTCACCATTCATCTTCATCTTGCACACT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCCCGACATCATCATGATTTCTTACCATCTTATGAGGCGCCATCCAGAAC 780
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnTrpAlaIle 280
 DB 781 TGGCTGGTACACATCATGCTCGACATTTGCCAATGCTTACCTTGAACACAGCCATC 840
 QY 281 AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaIleThrLeu 300
 DB 841 AACTCTTCTCTACTGCTTCATCAGCAAGCGTTCCGACCAATGAGCGCCAGCGCTC 960
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGCTTTCTTCAAGTCCAGAGCAACCTGTACAGTCTTACACCAATCATTAATCTTCC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 961 ATAAACAAGTAGCCCTGATCTCGCCGGAATCTACATGATCAAGATGCTGTGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1021 CAGTATGACAAATATGAAAAAAGTATTAAGATATCCCG 1059

RESULT 2
 ABK15562
 ID ABK15562 standard; cDNA; 1062 BP.
 XX
 AC ABK15562;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE cDNA encoding novel G-protein coupled receptor TGR8, version #1.
 XX
 KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
 KW cytostratic; antiinflammatory; antitumor; fetal brain;
 KW central nervous system disease; circulatory organ disorder; cancer;
 KW metabolic disease; immunological disease; gastrointestinal disease;
 KW gene therapy; transgenic animal; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers


```

FT CDS 1..1062
FT /tag= a
FT /product= "TGR8"
FT /note= "G-protein coupled receptor"
XX MO200194582-A1.
XX
XX 13-DEC-2001.
XX
XX 01-JUN-2001; 2001MO-JP004643.
XX
XX 02-JUN-2000; 2000JP-00170446.
XX 23-JUN-2000; 2000JP-00194926.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX Terao Y, Matsui H, Shintani Y;
XX
XX WPI; 2002-164317/21.
XX P-PSDB; AAU76416.
XX
XX Human fetal brain-originated G protein-coupled receptor protein TGR8 and
XX encoding DNA, for developing drugs to treat e.g. diseases of the central
XX nervous system or circulatory organs, cancer, and metabolic diseases.
XX
XX Claim 5; Page 94; 102pp; Japanese.
XX
XX The invention describes a human fetal brain-originated G protein-coupled
XX receptor protein, or its salt. The protein and encoded DNA are useful for
XX developing drugs to treat e.g. diseases of the central nervous system or
XX circulatory organs, cancer, metabolic diseases, immunological diseases
XX and gastrointestinal diseases. The invention also describes creation of a
XX probe for gene therapy and construction of a transgenic animal. This
XX sequence encodes the novel G-protein coupled receptor TGR8, described in
XX the method of the invention
XX
XX Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other:
XX
XX Alignment Scores:
XX Pred. No.: 6.83e-179 Length: 1062
XX Score: 1872.00 Matches: 353
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-09-995-225B-16 (1-353) x ABK15562 (1-1062)
XX
XX 1 MetGlnHiehrhfaiaiaiaaenSerleuSerProGly 20
XX 1 ATGAGACACACGACGCCACCTCGACGACCAAGCTGCTGGTGTCCTCCCGGC 60
XX
XX 21 SerAlaCyeglyleuGlyPheValProValValTyrTyrSerleuLeuCySleuGly 40
XX 61 TCGGCTGCGGCTTGGGTTTCGTGCCCGGTCTACACAGCTCTTGGCTGCTCGT 120
XX
XX 41 LeuProAlaValIleLeuThrValIleIleLeuSerGlnLeuValAlaGatGgInLys 60
XX 121 TTACACGACAAATATCTTGACAGATCATCTCTCCACAGTGGGCAAGAAACAGAG 180
XX
XX 61 SerSerTyrAsnTyrleuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
XX 181 TCTCTCTACACATATCTCTTGGCACTGCTGCTGCCACATCTTGGCTCTTTTCA 240
XX
XX 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
XX 241 GTGTTGTGACACTTCTGTTGGAAGATTTCATCTTGAAACAGAGCTCAGGTCC 300
XX
XX 101 AspLeuIleIleGluValLeuGluPheSerSerIleIleThrSerIleTrpIleThrVal 120
XX 301 GACAAATCATATAGAGGTGGAATTCATCATCCATCCACACTTCATATGATTAAGTGA 360
XX
XX 121 ProLeuThrIleAspArgTyrIleAlaValCysHiehrProLeuLysTyrIleThrValSer 140

```

```

DB 361 CGGTAAACATTACACAGGATATGCTGCTGCAACCCGCTCAAGATCACAGGCTCTCA 420
XX
XX 141 TyrProAlaArgThrArgIleValIleSerValTyrIleThrCysPheLeuThrSer 160
XX 421 TACCCAGCCCGCACCGGAAGCATTTGATGTTTACATCATCTGCTCTCCGACAGC 480
XX
XX 161 IleProTyrTyrTrpTrpProAsnIleTyrThrGluAspTyrIleSerThrSerValIle 180
XX 481 ATCCCTATTATACGTGGGCCCAACATCTGACATGAAAGACTACATCAGACCTTGTGCA 540
XX
XX 181 HisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
XX 541 CAGTCTCATCTGGATTCACATGCTTACACCGTCTACTGCTGCTGCTCATTCTTTC 600
XX
XX 201 IleLeuAsnSerIleIleValTyrIleValLeuArgArgLysSerAsnPheArgLeuArgLys 220
XX 601 ATCTTGACATCATCATTTGTGTCAGAGCTCAGGAGGAAAGCAATTTTGTCTCCGTGGC 660
XX
XX 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
XX 661 TACTCCACGGGGAAGACCAACCGCATCTGTTCACCATTAACCTCTTGGCACACTT 700
XX
XX 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValAspIleGlnAsnArg 260
XX 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTCTTATGGGGCGCCATCCAGAACCGC 780
XX
XX 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
XX 781 TGGCTGGTACATCATCATGTCGACATTCGCAACATGCTGACCTTGAACACAGCATC 840
XX
XX 281 AsnPhePheLeuTyrCysPheIleSerIleValArgPheArgThrMetAlaAlaThrLeu 300
XX 841 AACTTTTCTCTACGCTTTCATCAGCAAGCGGTTCGCAACCTGACCGCCAGCAGCTC 900
XX
XX 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
XX 901 AAGCTTTCTTCAAGTCCAGAAAGCAACCTGTACACATCTTACACCAATCACTTTTCC 960
XX
XX 321 IleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
XX 961 ATTAACAAGTAGACCCCTGATCTGCGCGCAACTCACATGATCAAGATGCTGGTGTAC 1020
XX
XX 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
XX 1021 CAGTATGACAAAATGAAAAACCTATMAAAGTATCCCG 1059
XX
XX
XX RESULT 3
XX ABK15563
XX ID ABK15563 standard; cDNA; 1062 BP.
XX
XX ABK15563;
XX
XX 08-MAY-2002 (first entry)
XX
XX cDNA encoding novel G-protein coupled receptor TGR8, version #2.
XX
XX G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
XX cytoskeletal; antiinflammatory; antitumor; fetal brain;
XX central nervous system disease; circulatory organ disorder; cancer;
XX metabolic disease; immunological disease; gastrointestinal disease;
XX gene therapy; transgenic animal; human; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..1062
XX /tag= a
XX /product= "TGR8"
XX /note= "G-protein coupled receptor"
XX
XX MO200194582-A1.

```


XX Chen R, Chu ZL, Dang HT, Lowitz KP, Pridie C;
 PI WPI; 2002-566565/60.
 DR P-PSDB; AB04075.
 XX Novel endogenous and non-endogenous versions of G protein-coupled
 PT receptor useful for identification of candidate compounds as receptor
 PS agonists or antagonists for use as therapeutic agents.
 XX Claim 31; Page 70; 84pp; English.
 CC The present invention provides the protein and coding sequences of
 CC several human G-protein coupled receptors (GPCRs). These can be used in
 CC the identification of candidate compounds as receptor agonists or inverse
 CC agonists having applicability as therapeutic agents. The present sequence
 CC is a GPCR coding sequence of the invention
 XX
 SO Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6, 83e-179 Length: 1062
 Score: 1872.00 Matches: 353
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-995-225b-16 (1-353) x AB04075 (1-1062)
 QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGGAGCACACGACGCCACCTCGACAGCCACAGCTCGCTGCTGGTGTCCCGCGC 60
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
 DB 61 TCGGCTGCGGCTTGGGTTTCTGCGCCGTGTCTACTACGCTCTTGGCTGCTGCTGCT 120
 QY 41 LeuProAlaAsnLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
 DB 121 TTAACGACAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGGCAGAGACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAlaAspIleLeuValLeuPhePheIle 80
 DB 181 TCTCTTACCAACATATCTTGTGACCTGCGCTGCGACATCTTGGCTCTTTTCATA 240
 QY 81 ValPheValAspPheLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 241 GTGTTGTGACCTTCTGTTGGAGATTTATCTTGAACTGAGATGCTCTCAAGTCCC 300
 QY 101 AspLysIleIleGluValLeuGlnPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 301 GACAAAGATCATAGAGGTGGAATCTCATCCATCCACATCCATATGATGATTACTGTA 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 361 CCCTTACCACTTACAGATATATCGCTGTGCACCCGCTCAAGTACCAACAGAGTCTCA 420
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCACCCGGAAAGTCAATTGATGTTTATCTACCTGCTCTTCTGACAGC 480
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTATATACGTGGGCCCAATCTGAGTGAAGACTAATCAGACACTGTCAT 540
 QY 181 HisValLeuIleThrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CAGGTCTCATCTGATCCACTGCTTCAACGCTTACCTGTGCTGCTGCTCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 601 ATCTTAACATCATCTGTGTACAACTCAGAGAGAGCAATTTTCTGCTCCGTGCGC 660

QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCCAGGGGAGAACACCGCATCTGTTCACATTAACCTTCATCTTTGCCACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTTATGGGGGCCCATCCGAAACCGC 780
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 781 TGGCTGTGACATCATATGTCACCATTCGCAACATGCTGACCTTGTGAACACAGCATC 840
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACCTTCTCTCTTACCTGCTTCAACACCAAGCGGTTCGCAACGTGGCAGCGCCAGCTC 900
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGGCTTTCTTCAAGTGCAGAGACACTGTACAGTTCTACACCAATCATTACTTTTCC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 961 ATAAACAAGTAGCCCTGATCTGCGCGGAACTCAACATGATCAAGATGCTGTGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1021 CAGTATGACAAATAATGAAAACTTATMAAAGTATCCCG 1059
 RESULT 5
 ACC44115
 ID ACC44115 standard; DNA; 1062 BP.
 XX ACC44115;
 AC
 XX
 DT 27-JUN-2003 (first entry)
 XX
 DE Human AXOR-57 coding sequence.
 XX
 XX de; antibacterial; antifungal; antiviral; antiparasitic; analgesic; gene;
 KM cytoactive; antidiabetic; anorectic; anabolic; antiepileptic; antagonist;
 KM antiparkinsonian; cardiac; hypotensive; hypertensive; nephrotoxic;
 KM osteopathic; antineural; antitumor; antiallergic; antimigraine; agonist;
 KM antiepileptic; nootropic; tranquilizer; neuroprotective; antidepressant;
 KM vaccine; gene therapy; G-protein coupled receptor; receptor; infection;
 KM pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; stroke;
 KM Parkinson's disease; osteoporosis; angina pectoris; ulcer; allergy;
 KM vomiting; schizophrenia; depression; dementia; Huntington's disease.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..1062
 FT CDS
 FT /*tag= a
 FT /product= "AXOR-57 protein"
 XX
 XX GB2365009-A.
 PD
 PD 13-FEB-2002.
 PF 11-APR-2001; 2001GB-00009018.
 XX
 XX 11-APR-2001; 2001GB-00009018.
 PR 11-APR-2001; 2001GB-00009018.
 XX
 XX (SMK) SMITHKLINE BEECHAM CORP.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 XX
 XX Elshourbagy N, Gattu M, Michalovich D, Shabon U,
 DR WPI; 2003-203569/20.
 DR P-PSDB; ABP98724.
 XX
 XX New G-protein coupled receptor, AXOR 57, for diagnosing and treating

PT New G protein-coupled receptor and encoded gene, useful in screening
 PT preventives or remedies for eating disorders or obesity.

XX Disclosure, Page 31-33; 48pp; Japanese.

XX The invention relates to a novel G protein-coupled receptor. The protein
 CC and its encoded gene are useful for screening preventives or remedies for
 CC eating disorders or obesity. The current sequence represents a G protein
 CC coupled receptor encoding sequence

XX Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6 83e-179	Length:	1062
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-995-225B-16 (1-353) x ACC71785 (1-1062)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpSerProGly 20
 DB 1 ATGGAGACACGCGACCGCCAGCTCGAGCCACAGCTCGCTGCTGTCCTCCCGC 60
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
 DB 61 TCGGCTGCGCGCTGGGTTTCGTGCGCGCTGCTACACAGCTCTTGCTGCTCGGT 120
 QY 41 LeuProAlaAsnLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnIle 60
 DB 121 TTACGACAAATATCTTGACAGTATCATCTCTCCAGCTGGGGCAGAAAGACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
 DB 181 TCCTCTCAACAATATCTTGACAGTATCATCTCTCCAGCTGGGGCAGAAAGACAGAG 240
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 241 GGTGTTGTGACATCTGCTGTTGGAAGATTCATCTTGAAATGACAGATGCTCAGGTC 300
 QY 101 AspPheIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 301 GACAAATATCAATAGTCTGGAATCTCATCCATCCACCTCCATGATGATTAAGTGA 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuTyrTyrHisThrValSer 140
 DB 361 CGGTAAACATTTACAGGTATATCGCTGCTGCGCACCGCTCAAGTACCAAGCTCTCA 420
 QY 141 TyrProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCACCCGGAAGTCAATGTGTTAATCATCACTGCTCCGACAGCAG 480
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTATTAATGAGTGGCCCAATCTGAGTGAAGACATCAATCAAGACCTCTGTGCAT 540
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CAGCTCTATCTGGATCCACTGCTTCAACGCTACTGCTGCTGCTGCTGCTGCTTCT 600
 QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgSerAsnPheArgLeuArgGly 220
 DB 601 ATCTTAACATCAATGATGTGTCAAGCTCAGAGGAGAAAGCAATTTTCTCTCCGTG 660
 QY 221 TyrSerThrGlySerThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrIle 240
 DB 661 TACTTCACGGGAGAAACCAACGCAATCTTGTCAACATTAACCTCCATCTTGCCACACT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCCCGGATCATGATGATTTCTTACCACTCTATGGGGGCGCCATCCAGAAACCG 780

QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
 DB 781 TGGCTGTACACATTCATGTCCTGACATTCGCAATCTGAGCTTCTGAAACAGAGCATC 840
 QY 281 AsnPhePheLeuTyrCysPheIleSerIleArgPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACTTCTTCTCTACCTGCTTTCATCAGCAAGCGTTCCGACCATGCGACCGCCAGCTC 900
 QY 301 LysAlaPhePheLeuCysGlnIysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGCTTCTCTCAAGCCAGCAAGCAACCTGTACAGTTCATCAACATCATTAATCTTTCC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysAlleIysMetLeuValTyr 340
 DB 961 ATTAACAAGTAGCCCTCGAATCTCGCGCAAACTCACTGCAATCAAGATGTGTGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1021 CAGTATGACAAATAATGAAAACCTATTAAGAATATCCCG 1059

RESULT 7

ABV73364
 ID ABV73364 standard; DNA; 1202 BP.

XX ABV73364;

XX 22-JAN-2003 (first entry)

XX Human TGR20 polypeptide encoding DNA.

XX G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;
 KW antilipemic; nootropic; neuroprotective; antianemic; antidiabetic; human;
 KW antiparkinsonian; antilipemic; TGR20; gene; ds.

OS Homo sapiens.

FH Key

FT CDS Location/Qualifiers

FT 68..1129

FT /tag= a

FT /product= "TGR20"

FT /note= "GPCR polypeptide"

PN MO200277001-A2.

XX 03-OCT-2002.

XX 08-MAR-2002; 2002MO-US007171.

XX 09-MAR-2001; 2001US-00802803.

XX 16-MAR-2001; 2001US-0276649P.

XX (TULA-) TULARIK INC.

XX Tian H, Zhao J, Chen J, Cutler G;

XX WPI; 2003-01881/01.

XX P-PSDB; ABB82499.

XX Claim 6; Page 63; 87pp; English.

CC The invention relates to G-protein coupled receptor (GPCR) polypeptides
 CC and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR83,
 CC TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
 CC useful for identifying compounds for treating a TGR-associated disorder,
 CC such as psoriasis, inflammatory bowel disease, hyperlipidemia,
 CC Parkinson's disease, Huntington's disease, anemia, immune and blood
 CC disorders, ulcerative colitis, Crohn's disease or spleen enlargement.

CC They are also useful for identifying cells such as kidney, liver,
 CC hypothalamus, colon, adipose, or spleen cells, for forensics and
 CC paternity determination, diagnosing diseases and examining signal
 CC transduction. The present sequence represents a human GR20 polypeptide
 CC encoding DNA

XX Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8 1e-179	Length:	1202
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-09-995-225b-16 (1-353) x ABV73364 (1-1202)

QY 1 MetGluHisThrHisAlaHisLeuAlaAAsnSerLeuSerTrpSerProGly 20
 DB 68 ATGAGACACACGACGCCGCTCGACAGCCACAGCTCGCTGTTGGTGTCCCGG 127
 QY 21 SerAlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCyLeuGly 40
 DB 128 TCGGCTGGGGCTTGGGTTTGGTGGCCCGGTCTACTACAGCTCTTGTGCTCGGT 187
 QY 41 LeuProAlaAsnLeuThrValIleIleLeuSerGlyLeuValAlaArgArgGly 60
 DB 188 TTACGACGAAATCTTGCACAGATCATCTCTCCAGCTGTGGCAAGACAGAA 247
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
 DB 248 TCCTCTACAACTATCTTGGGACCTGGCTGCGGACATCTGTCTCTTTTCATA 307
 QY 81 ValPheValAspPheLeuLeuGlyAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 308 GTGTTGTGACTCTCTGTGGAAAGTTTCACTTGAACATGACATGCTCAGGTCCC 367
 QY 101 AspLysIleIleGluValLeuGlyPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 368 GACAAATCATGAAGTGTGGAATCTCATCATCAACCTCCATAGATTTACTGTA 427
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 428 CCGTTAACCATGACGAGTATATCGCTGCTGCCACCCGCTCAAGACACAGGCTCA 487
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 488 TACCCAGCCCGACCCGAAAGTCATGTGAAGTTTACATCACCTGCTTCTGACCA 547
 QY 161 IleProTyrTyrTrpProAsnIleTrpThrGlyAspTyrIleSerThrSerValHis 180
 DB 548 ATCCCTATTAATCTGGTGGCCCAACATCTGACGAAAGTATACAGCACTCTGTGAT 607
 QY 181 HisValIleIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 608 CAGGCTTCATCTGATCCACGCTTCAACCTCTACCTGGTCCCTGCTCCATCTTCTTC 667
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgGlySerAsnPheArgLeuArgGly 220
 DB 668 ATCTTGAATCATATCTTGTACAGCTCAGAGAGAAAGCAATTTTGTCTCGTGC 727
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 728 TACTCACGGGGAAGACACCGCCCATCTTGTACCATTAACCTCATCTTGTCCACACTT 787
 QY 241 TrpAlaProArgGlyIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 788 TGGGGCCCCCGATATATATGATTTTACCACTCTATGGGGCCCGCATCAGAAACGC 847
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 848 TGGCTGTACATCATGTCCGACCTTGGCAACATGCTAGCCCTTCTGAACACAGCCATC 907

QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
 DB 908 AACTTCTTCTCTACTGCTTTCATTCAGCAAGCGGTTCGACCATGAGCCGCCACGCTC 967
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 968 AAGCTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTTACCAATCATATCTTCTCC 1027
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysAlaIleLeuValTyr 340
 DB 1028 ATACCAAGTACGCCCTGATCTCGCCGCAAACTCACACTGCATCAAGTGTGTGTA 1087
 QY 341 GlnTyrAspLysArgGlyLysProIleLysValSerPro 353
 DB 1088 CAGTATGACMAAAATGAAAACTTAAAGTATCCCG 1126

RESULT 8

ADP70587
 ID ADP70587 standard, DNA, 1776 BP.

XX ADP70587;

XX 12-FEB-2004 (first entry)

DE Orphan receptor ligand-related human protein gene SeqID210.

KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
 cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 GFPuv; Enhanced GFP; EGFP; human; gene; ds.

OS Homo sapiens.

PN WO2003071272-A1.

PD 28-AUG-2003.

PF 21-FEB-2003; 2003WO-JP001901.

PR 22-FEB-2002; 2002JP-00045728.

PR 23-OUL-2002; 2002JP-00213949.

PR 11-OCT-2002; 2002JP-00298237.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;

DR WPI; 2003-697654/66.

DR P-PSDB; ADP70485.

PT Transformation of cells with a fusion protein of an orphan receptor
 protein with a fluorescent protein useful for identification of ligands
 to the orphan receptor.

PS Example 4; SEQ ID NO 210; 594bp; Japanese.

CC This invention relates to a novel method of identifying ligands to an
 CC orphan receptor protein which comprises transforming cells with DNA
 CC encoding a fusion protein of the orphan receptor with a fluorescent
 CC protein, so that the fusion protein is expressed in the cells (or cell
 CC membranes isolated from them) and contacting the cells with the potential
 CC ligand to be tested. A suitable fluorescent protein for incorporation in
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
 CC identification of ligands binding to an orphan receptor protein.

XX Sequence 1776 BP; 471 A; 500 C; 358 G; 447 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1 39e-178	Length:	1776
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-09-995-225b-16 (1-353) x ADF70587 (1-1776)

QY 1 MetG|NH|eThri|sAla|H|sleu|Ala|A|asn|Ser|Ser|e|u|Ser|T|P|Ser|Pro|G|Y 20
 DB 1 ATGGAGCACACGACCGCCCTCGACGCCAACCTCGCTTGGTGGTCCCCGGC 60
 QY 21 Ser|A|C|y|e|g|y|e|u|g|y|P|he|V|a|I|y|T|y|Ser|e|u|e|u|C|y|e|u|g|Y 40
 DB 61 TCGGCGTGGCGGCTTGGGTTTGTGGCCGTGTCTACTACAGCCCTTGTGCTGCTCGGT 120
 QY 41 Leu|Pro|Ala|asn|I|le|u|Th|Val|I|e|I|e|u|Ser|G|I|e|u|V|a|I|a|a|g|G|I|n|Y 60
 DB 121 TTTACGACCAATATCTTGACAGATCATCTCTCCACGTGGTGGCAGACAGAG 180
 QY 61 Ser|Ser|T|y|a|n|T|y|le|u|e|u|A|l|e|u|A|A|A|a|s|p|I|e|u|V|a|I|e|u|P|he|P|I|e 80
 DB 181 TCTCTCTACACTATCTCTTGGCATCTGCTGCGCACATCTTGCTCTTTTCATA 240
 QY 81 Val|P|he|Val|A|s|P|he|Leu|e|u|G|I|u|A|s|P|he|I|e|u|A|n|Met|G|I|n|e|C|Pro|G|I|n|A|Pro 100
 DB 241 GTGTTGTGGACTTCCGTGGAAAGATTTCATCTTGAAACATGACAGATGCTCAGTCC 300
 QY 101 Asp|Ly|e|I|e|I|e|G|I|u|V|a|I|e|u|G|I|u|P|he|Ser|I|e|H|s|Th|Ser|I|e|T|P|I|e|Th|Val 120
 DB 301 GACAGATCATTAAGAGTGGAAATCTCATTCATCCACCTCCCATATGATTAATGATTA 360
 QY 121 Pro|Leu|Th|I|e|A|s|P|a|g|T|y|I|e|A|V|a|I|C|y|H|s|P|ro|e|u|y|T|y|H|s|Th|Val|Ser 140
 DB 361 CCGTTAACCTTACAGATATATCGCTGCTGCCACCCGCTCAAGTACACACAGGTTCA 420
 QY 141 T|y|P|ro|A|A|a|g|Th|A|g|y|V|a|I|e|V|a|S|e|r|V|a|I|y|T|I|e|Th|C|y|P|he|Leu|Th|Ser 160
 DB 421 TACCCAGCCCGCACCCGGAAAGTCAATGTAAAGTTAACTACACCTGCTTCGACACAGC 480
 QY 161 I|le|P|ro|T|y|T|T|P|T|P|P|ro|A|n|I|e|T|P|Th|G|I|u|A|s|P|T|y|I|e|S|e|Th|Ser|Val|H|s 180
 DB 481 ATCCCTATTAACGTGGGCCCAACATCTGACATGAAACATCAACACCTCTGTGCAT 540
 QY 181 H|s|V|a|I|e|u|I|e|T|P|I|e|H|s|C|y|P|he|Th|V|a|I|P|ro|C|y|S|e|r|I|e|P|he|P|he 200
 DB 541 CAGCTCTCATCTGGATCCAGCTGCTTCAACCTCTGCTGCTGCTCCTCATCTTCTTC 600
 QY 201 I|le|Leu|A|n|Ser|I|e|I|e|V|a|I|y|T|y|S|e|u|A|g|T|y|S|e|A|n|P|he|A|g|Leu|A|g|Y 220
 DB 601 ATCTTGAACATCAATGATGTGTAACAGCTCAGAGAGAAAGCAATTTTCCGTCGCGGC 660
 QY 221 T|y|S|e|Th|G|I|y|S|Th|Th|A|I|I|e|u|P|he|Th|I|e|Th|Ser|I|e|P|he|A|Th|Th|e|u 240
 DB 661 TACTCCACGGGGAAAGACACCGCATCTTGTTCACCATTCCTCCATCTTTGGCACACTT 720
 QY 241 T|y|P|A|P|ro|A|g|I|e|I|e|Met|I|e|u|T|y|H|s|I|e|u|T|y|G|I|A|P|ro|I|e|G|I|n|A|n|A|g 260
 DB 721 TGGGCCCCCGCATCATCATGATCTTTTACACCTCATATGGGGGCCCATCCAGAACCGC 780
 QY 261 T|y|P|e|u|V|a|I|s|I|e|Met|S|e|A|s|P|I|e|A|A|n|Met|e|u|A|l|e|u|e|u|A|n|Th|A|I|e 280
 DB 781 TGGCTGGTGACATCATGTCGACATTCGCAACATGCTACCCCTTGGAAACACAGCATC 840
 QY 281 A|s|n|P|he|P|he|Leu|T|y|C|y|P|he|I|e|S|e|T|y|A|g|P|he|A|g|Th|Met|A|A|A|a|I|a|Th|Th|e|u 300
 DB 841 AACTTCTTCTCTACAGCTTCATCAGCAACGGGTTCGCGACATGAGACGGCCAGCCTC 900
 QY 301 L|y|e|A|P|he|P|he|C|y|e|G|I|n|y|S|G|I|n|P|ro|V|a|I|G|I|P|he|T|y|Th|A|n|H|s|A|s|n|P|he|S|e|r 320
 DB 901 AAGGCTTTCTTCAAGTCCGGAAGCAACCTGTACAGTTCACCAATCATTAATTTTCC 960
 QY 321 I|le|Th|S|e|S|e|r|P|ro|T|y|I|e|S|e|r|P|ro|A|A|n|S|e|H|s|C|y|S|I|e|y|Met|e|u|V|a|I|y 340
 DB 961 ATAAACAAGTAGCCCTGATCTCGCCGGCAAACTCAACATGCATCAAGATGCTGTGTAC 1020

QY 341 G|I|n|T|y|A|s|P|y|A|s|n|G|I|y|S|P|ro|I|e|y|S|e|V|a|I|S|e|P|ro 353
 DB 1021 CAGTATGACAAATATGAAAACCTATATAAGTATCCCG 1059

RESULT 9
 ID AAS15731 standard; cDNA; 1343 BP.
 AC AAS15731;
 DT 14-FEB-2002 (first entry)
 DE DNA encoding chemokine receptor family related protein, NOV11.
 XX NOV, cytostatic; psoriasis; noctropic; neuroprotectant;
 XX cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
 XX haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;
 XX differentiation; proliferation; haematopoiesis; wound healing;
 XX angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
 XX haemophilia; allergy; pendred syndrome; skeletal dysplasia;
 XX ischemic injury; neuroepithelial disorder; hepatitis; heart failure;
 XX chemokine receptor; chromosome 1; ss.
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT 5'UTR 1 /*tag= a
 FT CDS 2..1063 /*tag= b
 FT /product= "Chemokine receptor related protein, NOV11"
 FT sig_peptide 2..142 /*tag= c
 FT mat_peptide 143..1060 /*tag= d
 FT /note= "Mature chemokine receptor related protein, NOV11"
 FT 3'UTR 1064..1343 /*tag= e
 XX
 XX MO200170978-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 20-MAR-2001; 2001MO-US009093.
 XX
 XX 20-MAR-2000; 2000US-0190768P.
 XX 20-MAR-2000; 2000US-0190835P.
 XX 22-MAR-2000; 2000US-0190972P.
 XX 22-MAR-2000; 2000US-0191199P.
 XX 24-MAR-2000; 2000US-0191947P.
 XX 28-MAR-2000; 2000US-0192657P.
 XX 28-MAR-2000; 2000US-0192654P.
 XX 28-MAR-2000; 2000US-0192665P.
 XX 28-MAR-2000; 2000US-0192884P.
 XX 29-MAR-2000; 2000US-0192836P.
 XX 31-MAR-2000; 2000US-0193843P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Taupier RJ, Majumder K, Spaderma SK, Smithson G, Mezes PS,
 XX Vernet CAM;
 XX WPI; 2001-639127/73.
 XX DR P-PSDB; AAU10068.
 XX
 XX Polypeptides and nucleic acids related to chloride channel, insulin-like
 XX growth factor family of proteins, useful for diagnosing and treating
 XX cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
 XX
 XX claim 9; Page 44; 15ip; English.
 XX
 XX The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
 XX polypeptides are useful for treating pathology associated with NOVX

CC polypeptide, determining the presence of or predisposition to a disease
CC associated with altered levels of NOX, identifying agents binding to
CC NOX and treatment of disorders associated with altered expression of
CC members of chloride channel-associated proteins e.g. cystic fibrosis and
CC congenital myotonia. NOX proteins are useful in treatment of disorders
CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
CC in one or more organs (e.g. haemophilia, anaemia), pendred syndrome,
CC skeletal dysplasia, disorders characterised by altered cell shape,
CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
CC treatment of disorders of vascular smooth muscle cell differentiation.
CC (e.g. heart failure, stroke). NOX nucleic acids and polypeptides are
CC useful to screen for molecules which inhibit or enhance NOX activity or
CC function and are useful as targets for the identifying small molecules,
CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
CC differentiation, haematopoiesis, wound healing and angiogenesis. NOX
CC sequences are also useful for: identifying a cell or tissue type in a
CC biological sample; amplifying DNA sequences from very small biological
CC samples e.g. hair or skin or body fluids and as primers and probes to
CC identify and/or clone NOX homologues. NOX proteins are useful
CC immunogens to generate antibodies to monitor protein levels and modulate
CC NOX activity. Cells comprising the nucleic acids are useful for
CC producing transgenic animals, for studying the function and/or activity
CC of NOX protein and identifying and/or evaluating modulators of NOX
CC protein activity. This sequence encodes NOV11 (located on chromosome 1)
CC related to the chemokine receptor family of proteins, one of 12 NOV
CC polypeptides described in the method of the invention

SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,39e-178 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 4 Gaps: 0

US-09-995-225b-16 (1-353) x AAS15731 (1-1343)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTyrSerProGly 20
DB 2 ATGGAGCAGACGACGCCCACTCGCAGCAACAGCTCGCTGTGGGTGCTCCCGCGC 61
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB 62 TCGGCTCGCGGCTTGGGTTTCGTGCCCGCTGTACTACAGCTCTTGCTGCTCGGT 121
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
DB 122 TTACCAAGCAAAATATCTTGACAGTGCATCTCTCCAGCTGGTGGCAAGAACAGAG 181
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
DB 182 TCCTCTCAACATTAATCTCTTGGCACTCGCTGTGCGCAATCTGGTCTCTTTTCTA 241
QY 81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 242 GTGTTTGGAGACTCTGTTGGAGATTTCACTTGAAACATGCAATCCCTCAGGTCCC 301
QY 101 AspLysIleIleGluValLeuGlnPheSerSerIleHisThrSerIleTrpIleThrVal 120
DB 302 GACAAAGATCATAGAGTCTGGAAATCTTCATCCACACCTCCATATGATTAATCTGA 361
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 362 CCGTTAACCATTTGACAGGTATATACCTGTGCCACCGCTCAAGTACCAACGAGTCTCA 421
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 422 TACCAAGCCCGACCCGGAAGAGTATGTATGTTTACATCACTGCTTCTGACCAAGC 481

QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGlnAspTyrIleSerThrSerValHis 180
DB 482 ATCCCTTATTAATGATGGGCCCAATCTGGACGTGAAGACTATACAGCACTCTGTGAT 541
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
DB 542 CACGTCCTCATCTGGATCCATCACTGTTCAACCGCTTACCTGTGCGCTCCATCTTCTTC 601
QY 201 IleLeuAsnSerIleIleValTyrTyrLeuArgArgLysSerAsnPheArgLeuArgGly 220
DB 602 ATCTTGAACTCATATGATGTGTACAGCTCAGAGAGAACCAATTTTCGTCCTGTGCG 661
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 662 TACTCCAGCGGGAACACACCGCATCTTGTTACCACTTACTCTCATCTTTGGCACACTT 721
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
DB 722 TGGGCCCCCGCATCATATGATTTCTTACCACTTATGGGGGCCCATCCAGAACCGC 781
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
DB 782 TGGCTGTACATCATATGTCGACATTCGCAATGCTAGGCTTCTGAAACAGACCATC 841
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
DB 842 AACTCTTCTCTCATCTGCTTCATCAGAACGCGTTCGCAACATGAGCGCGCACGCTC 901
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 902 AAGCTTTCTTCAAGTGCAGAGCAACACTGTACAGTTCTTACCAACCAATCACTTTTCC 961
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
DB 962 ATAAACAAGTACCCCTGATCTGCGCGCAAACTCAACATGATCAAGATGCTGTGTAC 1021
QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
DB 1022 CAGTATGACAAATAATGAAAACTATATAAGTATCCCG 1060

RESULT 10
ADJ87768
ID ADJ87768 standard; DNA; 1343 BP.
XX
AC ADJ87768;
DT 06-MAY-2004 (first entry)
XX
DB G-coupled protein receptor-related protein coding sequence #57.
XX novel protein; G-coupled protein receptor-related protein;
KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolemia;
KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
KW viral infections; bacterial infection; parasitic infection;
KW hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
KW Alzheimer's disease; tuberosus sclerosis; hypercalcaemia; cerebral palsy;
KW gene; ds.
XX
OS Unidentified.
XX
XX
PN MO2002102321-A2.
XX
PD 27-DEC-2002.
XX
PF 18-JUN-2002; 2002WC-US019522.
XX
PR 18-JUN-2001; 2001US-0298994P.
PR 18-JUN-2001; 2001US-0299134P.
PR 04-OCT-2001; 2001US-0097244P.
PR 06-JUN-2002; 2002US-00299134.
PR 07-JUN-2002; 2002US-00298994.

XX (CURA-) CURAGEN CORP.
 XX Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
 PI Pena CE, Spaderna SK, Zhong M;
 XX WPI: 2003-167441/16.
 DR P-PSDB; ADU87769.
 XX
 PT New MOLX polypeptides and polynucleotides, useful in gene therapy,
 PT particularly for treating or preventing e.g. cardiomyopathy,
 PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
 PT cancer, stroke or pain.
 XX
 PS Claim 8; SEQ ID NO 203; 378bp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of novel G-
 CC coupled protein receptor-related (MOL) proteins. The DNA and protein
 CC sequences of the invention are useful for treating or preventing a MOL-
 CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
 CC associated with cell signal processing and metabolic pathway modulation,
 CC or diabetes. The DNA and protein sequences are also useful for the
 CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
 CC pain, hypercholesterolemia, obesity, hypertension, Crohn's disease,
 CC systemic lupus erythematosus, viral infections, bacterial infections,
 CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or
 CC cerebral palsy. The present DNA sequence encodes a MOL protein of the
 CC invention.
 XX
 SO Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2.39e-178 Length: 1343
 Score: 1868.00 Matches: 352
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 99.79% Indels: 0
 DB: 10 Gaps: 0
 US-09-995-225B-16 (1-353) x ADJ87768 (1-1343)
 QY 1 MetGluNHiethrhfaaahleuuaalaasnserleuSerTPtPserProgly 20
 DB 2 ATGGAGCAGACGCGCCGACCTCGCAGCCAGCAGCGCTGGTGGTGGTCCCCGGC 61
 QY 21 SerAlaCyeglyeuGlPheValProValValTyTyrSerleuLeuCyseuGly 40
 DB 62 TCGGCTGGCGGCTTGGGTTTCGTGGCCGCTGCTACAGCCTCTTGGTCTGCTCGT 121
 QY 41 LeuProAlaAsnIleuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
 DB 122 TTACGAGCAAAATATCTTGACAGATGATCATCTCTCCAGCTGGTGCAAAAGAGAG 181
 QY 61 SerSerTyrrasnTyrrleuAlaAlaAlaAlaAspIleleuValIleuPhePheIle 80
 DB 182 TCTCTCCACACATATCTTGGGACCTGCTGCGCCAGCATCTTGGTCTCTTTTCATA 241
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 242 GTCTTTGTGACCTTCTGTTGGAAAGATTTCATCTTGAAACATGCAAGATGCTCAGTCC 301
 QY 101 AspLysIleIleGluValIleuGluPheSerSerIleIleThrSerIleTrrIleThrVal 120
 DB 302 GACAAATCATATAAGGTGAGATTCATCATCATCACACCTCCATATGATTAACGTGA 361
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 362 CGGTAAACATTAACAGATATATCATCTGTGCACCCGGCTCAAGATCAACACAGGTCTCA 421
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 422 TACCGAGCCGCAACCGGAAGATCATTTGTAAGTTTAACTACATCACTGCTTCTGACACG 481

QY 161 IleProTyrTyrTrpTrpProAlaAsnIleTrrPThrGluAspTyrIleSerThrSerValHis 180
 DB 482 ATCCCTTATTTAGTGGGCGCCCAACATCTGAGCTGAAGACTACATCACTGCTGGCAT 541
 QY 181 HisValIleuIleTrrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 542 CAGTCTCTATCTGGATCCATGCTTCCACCGCTTACCTGCTGCTCCCTCCATCTTCTTC 601
 QY 201 IleLeuAsnSerIleIleValTyrIysLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 602 ATCTTAACATCATCATCTGTGTACAAAGCTCAGAGAGAGAGCAATTTGCTCCGTGAC 661
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 662 TACTCCACGGGAGAGCCACCGCATCTTGTTCACCATTAACCTTCATCTTGGCACACTT 721
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValaProIleGlnAsnArg 260
 DB 722 TGGGCCCCCGCATCATCATGATTTCTTTCACACTTATGGGGCGCCATCCAGAACCCG 781
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 782 TGGCTGTACATCATGTGCCAGATGCCAATGCTGACCTTCTGAACACAGCCATC 841
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysValArgPheArgThrMetAlaAlaThrLeu 300
 DB 842 AACTTCTTCTTACGCTTTCATCAGCAAGCGGTTCGGACCAATGGCACCGCCAGCATC 901
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValAlaGlnPheTyrThrAsnHisAsnPheSer 320
 DB 902 AAGCTTTTCTTCAAGGCCAGAGCAACCTGACAGTTTACACCATATTAATCTTTC 961
 QY 321 IleThrSerSerProTrrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 962 ATAAACAAGTAGCCCGCTGATCTCGCCGCAAACTCACATGATCAAGATGTGCTGATC 1021
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1022 CAGTATGACMAAAATGGAACCTTATMAAAGTATCCCG 1060
 RESULT 11
 ADI79324
 ID ADI79324 standard; DNM: 1343 BP.
 XX
 AC ADI79324;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE NOV11 coding sequence, SEQ ID 21.
 XX
 KW Cytostatic; Immunosuppressive; Anti-allergic; Antimicrobial; Vasotropic;
 KW Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
 KW Ophthalmological; Antipneumatic; Neuroprotective; Nootropic;
 KW Antiartherosclerotic; Hypotensive; Cardiant; Cerebroprotective;
 KW Gene Therapy; NOX; human; cancer; myelogenous leukemia;
 KW congenital neonatal autoimmune chromocyclopaenia; immunological disorder;
 KW allergy; infection; asthma; lung disease; reproductive disorder;
 KW haemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;
 KW diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;
 KW hypertension; stroke; heart failure; chromosome 1; NOV11;
 KW chemokine receptor; gene; de.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 2..1063
 FT /tag= a
 FT /product= "NOV11"
 XX
 XX MO2004009635-A2.
 XX
 PD 29-JAN-2004.

XX 04-OCT-2001; 2001WO-US031292.
PF
XX
PR 20-MAR-2001; 2001US-00913432.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS,
PI Verneir CM;
XX
DR WPI; 2004-123380/12.
DR P-PSDB; ADI79325.
XX
PS Claim 8; Page 43-44; 158pp; English.
XX
CC The present invention relates to novel NOVX proteins and their coding
CC sequences (ADI79304-ADI79327). The sequences are useful for the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease associated with the protein, or for diagnosing and treating
CC disorders associated with the NOVX protein, such as cancer, myelogenous
CC leukaemia, congenital neonatal autoimmune thrombocytopenia,
CC immunological disorders, allergy and infection, asthma, lung diseases,
CC reproductive disorders, male and female reproductive diseases,
CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,
CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
CC hypertension, stroke and heart failure. NOV1 represents a new member of
CC the chemokine receptor family and the gene is located on chromosome 1.
XX
SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,39e-178 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 12 Gaps: 0
US-09-995-225b-16 (1-353) x ADI79324 (1-1343)
QY 1 MetGLuHsIeThrHsIaHsIleuA1a1aenSerSerleuSerTrpTrpSerProGly 20
DB 2 ATGGAGCAACGACGCCACCTCGGACGCAACAGCTCGCTTGGTGGTCCCGCCG 61
QY 21 SerAlaCyGlyLeuGlyPheValProValValTYTYTYSerleuLeuGlyCysLeuGly 40
DB 62 TCGGCTGCGGCTTGGGTTTCGGCCGCGTCTACTACAGCTCTGCTGCTCGT 121
QY 41 LeuProAlaAsnIleuLeuThrValIleIleleuSerGlnLeuValAlaArgArgGlnlys 60
DB 122 TTACCGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGAGACAGAAG 181
QY 61 SerSerTYrAsnTYrleuLeuAlaLeuAla1a1aenSerSerleuValleuPhePheIle 80
DB 182 TCCTCTTACAACTATCTCTTGGCACTCGCTGCGGACATTTTGCTCTTTTCAATA 241
QY 81 ValPheValAspPheLeuLeuGlnAspPheIleleuAsnMetGlnMetProGlnValPro 100
DB 242 GTGTTTGGACCTCTGTTGGAAATTTTCATCTTGAAATGCAATGCAATGCTCAGTCCC 301
QY 101 AspLysIleIleGlnValleuGlnPheSerSerIleHsIeThrSerIleTrpIleThrVal 120
DB 302 GACAGATCATTAAGAGTGTGAATTCATCATCCACACCTTCATATGATTAATCTGTA 361
QY 121 ProLeuThrIleAspArgTYrIleAlaValCYSHsIProLeuLysTYrHsIeThrValSer 140
DB 362 CCGTTAACCATTTGACAGGATATCTCTGCTCCACCCGCTCAAGTACCAACGGTCTCA 421
QY 141 TYrProAlaArgThrArgLysValIleValSerValTYrIleThrCYSPheLeuThrSer 160
DB 422 TACCCAGCCCGACCCCGAAAGTCAATGTAGTGTATCATCACCTGCTTCTGACGACG 481
QY 161 IleProTYrTYrTrpTrpProAsnIleTrpThrGlnAspTYrIleSerThrSerValHis 180

DB 482 ATCCCTATTACGTGGGCCCAACATCTGATGTAAGACTATACGACCTCTGTGAT 541
QY 181 HisValleuIleTrpIleHsIeCyPheThrValTYrleuValProCySerIlePhePhe 200
DB 542 CAGTCTCTATCTGATCCATCGCTTCACCGCTCACTGAGTCCCTGCTCCATCTTCTTC 601
QY 201 IleleuAsnSerIleIleValTYrLysleuArgLysSerAsnPheArgLeuArgGly 220
DB 602 ATCTGAATCATATGTTGTACAGGCTCAGAGAGAGAGCAATTTTCGTCTCCGTGGC 661
QY 221 TYrSerThrglyLysThrThrAlaIleleuPheThrIleThrSerIlePheAlaThrleu 240
DB 662 TACTCCAGCGGAGAACACACCGCATCTTGTTCACCATTAATCCATCTTGGCCACACTT 721
QY 241 TrpAlaProArgIleIleMetIleleuTYrHsIeLysTYrGlyAlaProIleGlnAsnArg 260
DB 722 TGGGCCCCCGCATATCATGATTTTACCACTCTATGGGCGCCCATCCAGAACCGC 781
QY 261 TrpLeuValHsIeMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
DB 782 TGGCTGTACACATCATGTCCGACATTCGCAACATGTAACCTTGAACAGCCATC 841
QY 281 AsnPhePheLeuTYrCYSPheAlaIleSerLysArgPheArgThrMetAlaAlaThrleu 300
DB 842 AACTCTTCTCTACTGCTTATCATGACGAAGCGGTTCCGACACATGAGCGCCGACGCTC 901
QY 301 LysAlaPhePheLysCYGlnLysGlnProValGlnPheTYrThrAsnHsIAsnPheSer 320
DB 902 AAGCTTTCTTCAAGTGCAGAGCAAGCACTGTATACAGTTCTACACCATCATTAATTTCC 961
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHsIeLysMetLeuValTYr 340
DB 962 ATAAACATATACCCCTGATCTCGCGCAAACTCACACATGATCAAGATGCTGTATAC 1021
QY 341 GlnTYrAspLysAsnGlyLysProIleLysValSerPro 353
DB 1022 CAGTATGACAAATGGAACCTATTAAGTATCCCG 1060
RESULT 12
AD056003
ID AD056003 standard; cDNA; 1343 BP.
XX
AC AD056003;
XX
DT 15-JUL-2004 (first entry)
XX
DE DNA encoding human NOV1.
XX
KW human; gene; ss; cancer; obesity; diabetic nephropathy;
XX acute pancreatitis; stroke; multiple sclerosis.
XX
OS Homo sapiens.
XX
PN US2004058862-A1.
XX
PD 25-MAR-2004.
XX
PF 18-SEP-2002; 2002US-00246583.
XX
PR 18-SEP-2002; 2002US-00246583.
XX
PA (MAJU/) MAJUMDER K.
XX
PI Majumder K;
XX
DR WPI; 2004-26835/25.
DR P-PSDB; AD056004.
XX
PT Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic
XX nephropathy, acute pancreatitis, strokes and multiple sclerosis.
PS Disclosure; Page 28; 87pp; English.

XX The invention relates to novel isolated NOVX nucleic acids and encoded polypeptides. The nucleic acids, polypeptides and antibodies raised against the polypeptides are useful for preventing or treating diseases associated with aberrant NOVX expression or activity e.g., cancer, obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple sclerosis. The present sequence represents a NOVX nucleic acid of the invention.

XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Alignment Scores:

Prod. No.:	2,39e-178	Length:	1343
Score:	1868.00	Matches:	352
Percent Similarity:	99.72%	Conservative:	0
Best Local Similarity:	99.72%	Mismatches:	1
Query Match:	99.79%	Indels:	0
DB:	12	Gaps:	0

US-09-995-225b-16 (1-353) X ADO56003 (1-1343)

QY 1 MetGluHisThrHisAlaHisAlaHisSerLeuSerTrpSerProGly 20
 DB 2 ATGAGACACACGACCGCCACCTCGACCAAGCTCGTCTTGTGTCCTCCCGC 61
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuGly 40
 DB 62 TCGGCGCGCGCTGGTTCGTGCGCGGTCTACTACAGCCCTTGTGCTCGT 121
 QY 41 LeuProAlaSerLeuThrValIleIleLeuSerGlnLeuValAlaArgArgLys 60
 DB 122 TTACACCAATATCTTGACAGTATATCTCTCCAGCTGGTGGCAAGACAGAG 181
 QY 61 SerSerTyrAntTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
 DB 182 TCTCTCAACAATATCTCTTGACAGCTCGTGGCGACACCTTGGTCTTTTCATA 241
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 242 GTTGTGTGACCTTCTGTTGAGAGATTATCTTGAACATGACAGATCGCTGCTTTCATA 301
 QY 101 AspLeuIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 302 GACAAAGTCTAAGAGTCTGGAATTCATTCATCCACACCTCCATATGATTAAGTGA 361
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 362 CCGTTAACATATGACAGTATATCACTGTCGACACCGCTCAAGTACCAACGGTCTCA 421
 QY 141 TyrProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 422 TACCCAGCCCGCACCCGGAAGTCAATGTAAGTGTATACATCACTGCTCTCGACAGC 481
 QY 161 IleProTyrTyrTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 DB 482 ATCCCTATTAATCTGAGTGGCCCAACATTTGACAGAACTACATCAACACTCTGTCAT 541
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 542 CACGTCCTCATCTGAGTCCACTGCTTCACTGAGTGGCCCTGCTCATCTTCTTC 601
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
 DB 602 ATCTTAACATCAATCATGTGTACAGCTCAGGAGGAGCAATTTCTCTCCCGTGC 661
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 662 TACTCAAGGAGAGACACCGCATTTTTCACCTTACCTCCATCTTTGCCACACTT 721
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 722 TGGGCCCCCGCATCATGATCTTTTACACCTATGAGGGGCCCATCCAGAACCGC 781

QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 782 TGGCTGTACACATCAATGTCGACATTCGCAACATCTAGCCTTCTGAAACAGCCATC 841
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
 DB 842 AACTTCTCTCTACAGCTTCTCATACGCAAGCGGTTCCGACATGACAGCCGACGCTC 901
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 902 AAGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTCTACACCATCAATCAATCTTTC 961
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
 DB 962 ATAAACAGTACCCCTCGAGTCTGCGCGCAAACTCACTGCAATCAAGATGTGTATC 1021
 QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
 DB 1022 CAGTATGACAAATAATGAAATCCTATTAAGTATCCCG 1060

RESULT 13

ADO28955 standard; cDNA; 2273 BP.

XX ADO28955;
 XX 29-JUL-2004 (first entry)
 XX Human novel GPCR PCR polynucleotide, SEQ ID NO:54.

XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
 KW cytostatic; antiinflammatory; vasotropic; antianginal; antianthemic;
 KW CNS; central nervous system; respiratory; antidiarrhetic; antidiabetic;
 KW viricide; hepatotropic; antibacterial; antianemic; antidiarrhetic;
 KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KW gene; ss.

XX Homo sapiens.

XX WO2004/040000-A2.

XX 13-MAY-2004.

XX 09-SEP-2003; 2003WO-US028226.

XX 09-SEP-2002; 2002US-0409303P.

XX 09-APR-2003; 2003US-0461329P.

XX (PRIM-) PRIMAL INC.

XX Galenaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;

XX Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;

XX WPI; 2004-390329/36.

XX DR P-PSDB; ADO28954.

XX Novel mammalian G protein coupled receptors, useful for identifying

XX PT compounds that modulates diagnosing and treating disease condition

XX PT associated with GPCR dysfunction e.g. autoimmune diseases, angina

XX PT pectoris, Parkinson's disease.

XX Claim 13; SEQ ID NO 54; 542pp; English.

XX The invention relates to human and mouse G protein-coupled receptors

(GPCRs) and nucleic acids encoding them. The invention also relates to sequences at least 90% identical to the GPCR proteins and nucleic acids of the invention; methods of treating, preventing or diagnosing diseases associated with GPCRs of the invention; methods of screening for compounds useful in the treatment of GPCR-related diseases; a transgenic mouse comprising a GPCR gene of the invention; a mouse comprising a mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived from the transgenic mice; kits comprising several GPCR genes; each of which has a mutation in a different GPCR gene of the invention; and kits comprising probes which hybridize to GPCR polynucleotides of the invention. The invention further discloses variants of the GPCR polypeptides and vectors comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may be used in the diagnosis, treatment or prevention of a wide variety of diseases including neurological disorders (e.g., Alzheimer's disease, depression, diabetic neuropathy, Parkinson's disease or schizophrenia), disorders of the adrenal gland; disorders of the colon or intestine (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or myocardial infarction); muscular disorders; blood disorders (e.g., anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g., obesity, enzyme deficiency-related diseases or vitamin deficiency-related diseases); and disorders of the kidney, liver, lung, breast, ovary, uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and thyroid (e.g., cancers). The present sequence represents a GPCR-encoding nucleic acid of the invention. Note: The full sequence data for this patent did not form part of the printed specification; those sequences not shown were obtained in electronic format directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences.

80 Sequence 2273 BP; 503 A; 691 C; 549 G; 530 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,95e-178 Length: 2273
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 12 Gaps: 0

US-09-995-225b-16 (1-353) x ADO26955 (1-2273)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpSerProGly 20
DB 453 ATGGAGACACGACGCCCACTCGACGCAACAGCTCGCTGTGGTGGTCCCGGCG 512
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuGlyCysLeuGly 40
DB 513 TCGGCTGGGGGCTTGGGTTTCGTCGCCCGCTGCTACTACAGCCCTCTGCTGCTCGGT 572
QY 41 LeuProAlaAsnLleLeuThrValLleLleLeuSerGlnLeuValAlaArgArgGlnLys 60
DB 573 TTACACGACAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAG 632
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspLleLeuValLeuPhePheLe 80
DB 633 TCTCTCTTAACATATCTCTTGACACTCGCTGCGCACTTGGTCTCTTTTTCATA 692
QY 81 ValPheValAspPheLeuGlnLysAspPheLleLeuAsnMetGlnMetProGlnValPro 100
DB 693 GTGTTGTGAGACTCTCTGTGGAGATTTCATCTTGAACATGACAGATCCAGGTCCC 752
QY 101 AspLysLleLleGluValLeuGlnLysSerSerLleHisThrSerLleTyrLleThrVal 120
DB 753 GACACGATCATAGAGTCTGGAATTCATCATCCACACCTCATATGATTAAGTGA 812
QY 121 ProLeuThrLleAspArgTyrLleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 813 CCGTTAACCATTTGACAGGTATATCATCTCTGCGCAACCGCTCAAGTACCAACGCTCA 872
QY 141 TyrProAlaArgThrArgLysValLleValSerValTyrLleThrCysPheLeuThrSer 160

DB 873 TACCCAGCCGCGACCCGGAAGTCATTGTAAGTGTATACATCATCCTGTTCCGACACG 932
QY 161 IleProTyrTyrThrProAsnLleTyrPThrGlnAspTyrLleSerThrSerValHis 180
DB 933 ATCCCTTATTCATGTTGGGGCCCAACATGTGACTGAAGACTATATACATGACACCTCTGTGCA 992
QY 181 HisValLeuLleTyrLleHisCysPheThrValTyrLeuValProCysSerLlePhePhe 200
DB 993 CAGTCTCATCTGATGCATCATGCTTACCCGCTACCTGTTGCTCCATCTTCTTC 1052
QY 201 IleLeuAsnSerLleLleValTyrTyrLeuAlaGlyLysSerAsnPheArgLeuArgLys 220
DB 1053 ATCTTGACTCATATATGTGTGACAGCTCAGAGAGAAAGCAATTTTGCTCCGTGCG 1112
QY 221 TyrSerThrGlyLysThrThrAlaLleLeuPheThrLleThrSerLlePheAlaThrLeu 240
DB 1113 TACTCAGAGGGAAGAACACCGCCATCTTGTTACCATTTACCTTCATCTTGGCACACTT 1172
QY 241 TrpAlaProArgLleLleMetLleLeuTyrHisLeuTyrGlyAlaProLleGlnAsnArg 260
DB 1173 TGGGCCCCCGCATCATCATGATTCCTTACACCTCTATGGGGCGGCCCATCCAGAACCGC 1232
QY 261 TrpLeuValHisLleMetSerAspLleAlaAsnMetLeuAlaLeuLeuAsnThrAlaLle 280
DB 1233 TGGCTGGTACATATATGTCCGACATTCGCAACATGTGAGCCCTTGGAAACAGCCATC 1292
QY 281 AsnPheLeuTyrCysPheLleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
DB 1293 AACTCTTCTCTACTGCTTATCATCAGCAAGCGGTTCGACACATGGACGCCGCCACCGCTC 1352
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 1353 AAGGCTTCTTCAAGTGCACAAAGCAACCTGTACAGTTCTTACACCAATCATTAACCTTTTCC 1412
QY 321 IleThrSerSerProTyrLleSerProAlaAsnSerHisCysLleLysMetLeuValTyr 340
DB 1413 ATTAACAAGTACCCCTGGATCTGCGCGCAAACTCAACATGCATCAAGATGCTGTGTAC 1472
QY 341 GlnTyrAspLysAsnGlyLysProLleLysValSerPro 353
DB 1473 CAGTATGACAAATAATGAAAACTATATAAAGTATCCCGC 1511

RESULT 14
AAS15730
ID AAS15730 standard; cDNA; 1119 BP.
XX
AC AAS15730;
XX
DT 14-FEB-2002 (first entry)
XX
DE DNA encoding chemokine receptor family related protein, NOY10.
XX
KW NOV, cytosstatic; porlasis; nootropic; neuroprotectant;
KW cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
KW haemostatic; atherosclerosis; gene therapy; neurogenesis; mobility;
KW differentiation; proliferation; haematopoiesis; wound healing;
KW angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
KW hemophilia; allergy; Pendred syndrome; skeletal dysplasia;
KW ischemic injury; neuroepithelial disorder; hepatitis; heart failure;
KW chemokine receptor; chromosome 1; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1119
FT /tag= a
FT /product= "Chemokine receptor related protein, NOY10"
FT sig_peptide 1..141
FT /*tag= b
FT 142..1116
FT /*tag= c
FT /note= "Mature chemokine receptor related protein, NOY10"
XX

PN MO200170978-A2.
 XX 27-SEP-2001.
 XX 20-MAR-2001; 2001MO-US009093.
 XX 20-MAR-2000; 2000US-0190768P.
 PR 20-MAR-2000; 2000US-0190835P.
 PR 22-MAR-2000; 2000US-0190972P.
 PR 22-MAR-2000; 2000US-0191199P.
 PR 24-MAR-2000; 2000US-0191947P.
 PR 28-MAR-2000; 2000US-0192657P.
 PR 28-MAR-2000; 2000US-0192664P.
 PR 28-MAR-2000; 2000US-0192655P.
 PR 28-MAR-2000; 2000US-0192844P.
 PR 29-MAR-2000; 2000US-0192836P.
 PR 31-MAR-2000; 2000US-0193843P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;
 PI Vernet CAM;
 DR WPI; 2001-639127/73.
 DR P-PSDB; AAU10067.
 XX
 PT Polypeptides and nucleic acids related to chloride channel, insulin-like
 PT growth factor family of proteins, useful for diagnosing and treating
 PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
 XX
 PS Claim 9; Page 42; 151pp; English.
 XX
 CC The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX
 CC polypeptides are useful for treating pathology associated with NOVX
 CC polypeptides, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX, identifying agents binding to
 CC NOVX and treatment of disorders associated with altered expression of
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and
 CC congenital myotonia. NOVX proteins are useful in treatment of disorders
 CC including porphyria, cancer, diabetes, metabolic disorders of pancreas,
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
 CC in one or more organs (e.g. haemophilia, anaemia), Pendered syndrome,
 CC skeletal dysplasias, disorders characterised by altered cell shape,
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
 CC treatment of disorders of vascular smooth muscle cell differentiation,
 CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
 CC useful to screen for molecules which inhibit or enhance NOVX activity or
 CC function and are useful as targets for the identifying small molecules,
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
 CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV
 CC sequences are also useful for: identifying a cell or tissue type in a
 CC biological sample; amplifying DNA sequences from very small biological
 CC samples e.g. hair or skin or body fluids and as primers and probes to
 CC identify and/or clone NOVX homologues. NOVX proteins are useful
 CC immunogens to generate antibodies to monitor protein levels and modulate
 CC NOVX activity. Cells comprising the nucleic acids are useful for
 CC producing transgenic animals, for studying the function and/or activity
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX
 CC protein activity. This sequence encodes NOV10 (located on chromosome 1)
 CC related to the chemokine receptor family of proteins, one of 12 NOV
 CC polypeptides described in the method of the invention
 XX
 SQ Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,39e-177 Length: 1119
 Score: 1857.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.20% Indels: 0
 DB: 4 Gaps: 0

US-09-995-225b-16 (1-353) x AAS15730 (1-1119)
 QY 1 MetGLNHISThrHISAlaHISLeuAlaAlaSerSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGAGACACACGACCCACCTCGACGCAACAGCTCGTCTTGCGTCCCCGCG 60
 QY 21 SerAlaCyGAGlyLeuLeuIlePheValProValIleTyrTyrSerLeuLeuLeuGly 40
 DB 61 TCGGCTGGGCTGGGTTTCGCGGTGTACTACAGCTCTTGCTGCTCGGT 120
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGln 60
 DB 121 TTACGACGAATATCTTGACAGATCATCTCTCCAGCTGGTGGCAAGACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
 DB 181 TCCTCTACAACTATCTTGGCACTCGGTGCGGACACTCTTGCTCTTTTCTATA 240
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 160
 DB 241 GTGTTGTGACTTCTGTGTGAAGATTTCATCTTGAACATGCAATGCTTCAGTCCC 300
 QY 101 AspValIleIleGluValLeuGluPheSerSerIleIleThrSerIleTrpIleThrVal 120
 DB 301 GACAAATCATAGAAATGCTGGATTCATCATCATCACAACCTCATATGATTTACTGTA 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLeuTyrHisThrValSer 140
 DB 361 CCGTTAACATGACAGATATATCGCTGTGCGCACCCGCTCAAGTACACAGGTCTCA 420
 QY 141 TyrProAlaArgThrArgLeuValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGACCCGGAAGTCATGTGAAGTTTACATCCCTGCTTCGACACAG 480
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpGlnAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTTATTAATCGGTGGCCCAACATCGATGGAAGATCATCAGCACTGTGCAAT 540
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CACGTCTCATCTGATGATCCACGCTTCACCGTCTACCTGTGCGCTCCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrIleValAspArgLeuSerAsnPheArgLeuAspGly 220
 DB 601 ACTTGAATCATCATCTTGTGACAGCTCAGAGGAAGACAAATTTGCTCCGTGG 660
 QY 221 TyrSerThrGlyLeuThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCAGCGGGAAGACCAACCCCATCTTGTTACCATTAACCTCTTGCCACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTGATGGGCGCCCATCCGAACCCG 780
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
 DB 781 TGGCTGTGACATCATCTGTCACATGTCACATGTCAGTCCCTTGTGAACACAGCATTC 840
 QY 281 AsnPhePheLeuTyrCysPheIleSerIleAspPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACTTCTTCTCTACAGCTTCACTCATCAGCAAGCGGTTCGCAACATGCGACCGCAGCTC 960
 QY 301 LysAlaPhePheLeuCysGlnLeuGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGGCTTCTTCAAGGCGCAGAAAGCAACCTGTACAGTTCTACACCAATCATTAATCTTTC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLeuMetLeuValTyr 340
 DB 961 ATTAACAAGTAGAGCCCTGGATCTCGCGGCAAACTCACTGATCAAGATGCTGGTGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLys 350
 DB 1021 CAGTATGACAAATAATGGAATAACTTATATAA 1050

RESULT 15
 ID ADJ87766 standard; DNA; 1119 BP.
 AC ADJ87766;
 DT 06-MAY-2004 (first entry)
 DE G-coupled protein receptor-related protein coding sequence #56.
 XX
 KM novel protein; G-coupled protein receptor-related protein;
 KM cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
 KM metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
 KM Huntington's disease; epilepsy; anxiety; pain; hypercholesterolemia;
 KM obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
 KM viral infections; bacterial infection; parasitic infection;
 KM hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
 KM Alzheimer's disease; tuberos sclerosis; hypercalcaemia; cerebral palsy;
 KM gene; ds.
 XX
 OS Unidentified.
 PN MO2002102321-A2.
 PD 27-DEC-2002.
 XX
 PF 18-JUN-2002; 2002MO-US019522.
 XX
 PR 18-JUN-2001; 2001US-0298994P.
 PR 18-JUN-2001; 2001US-0299134P.
 PR 04-OCT-2001; 2001US-00972446.
 PR 06-JUN-2002; 2002US-00299134.
 PR 07-JUN-2002; 2002US-00298994.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
 PI Pena CE, Spaderna SK, Zhong M;
 DR WPI; 2003-167441/16.
 XX
 DR P-PSDB; ADJ87767.
 XX
 PT New MOX polypeptides and polynucleotides, useful in gene therapy,
 PT particularly for treating or preventing e.g. cardiomyopathy,
 PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
 PT cancer, stroke or pain.
 XX
 PS Claim 8; SEQ ID NO 201; 378bp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of novel G-
 CC coupled protein receptor-related (MOX) proteins. The DNA and protein
 CC sequences of the invention are useful for treating or preventing a MOX-
 CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
 CC associated with cell signal processing and metabolic pathway modulation,
 CC or diabetes. The DNA and protein sequences are also useful for the
 CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
 CC pain, hypercholesterolemia, obesity, hypertension, Crohn's disease,
 CC systemic lupus erythematosus, viral infections, bacterial infections,
 CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, tuberos sclerosis, hypercalcaemia, or
 CC cerebral palsy. The present DNA sequence encodes a MOX protein of the
 CC invention.
 XX
 SQ Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,39e-177 Length: 1119
 Score: 1857.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.20% Indels: 0
 DB: 10 Gaps: 0

US-09-995-225B-16 (1-353) x ADJ87766 (1-1119)
 QY 1 MetGluHisThrHisAlaHisIleuAlaIleuAsnSerSerIleuSerTrpTrpSerProGly 20
 DB 1 ATGAGACACAGCAGCCGCCCTCGCAGCAACAGCTCGCTCTTGGGGTCCCCCGGC 60
 QY 21 SerAlaCyseGlyLeuGlyPheValProValValTYrTYrSerIleuLeuCyseGly 40
 DB 61 TCGGCCTGGCGCTTGGGTTTGGTGGCCCGGTCTGCTACAGCTTGTGCTCGCT 120
 QY 41 LeuProAlaAsnIleuLeuThrValIleIleuSerGlnLeuValAlaArgArgGlnlys 60
 DB 121 TTACACAGCAAAATATCTTACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAA 180
 QY 61 SerSerTYrAsnTYrLeuLeuAlaIleuAlaAlaAspIleuValIleuValPhePheIle 80
 DB 181 TCTCTCTACACTATCTTGTGGCACTCGCTGGCGCATCTTGGCTCTCTTTTTCATA 240
 QY 81 ValPheValAspPheLeuLeuGlyAspPheIleuAsnMetGlnMetProGlnValPro 100
 DB 241 GTGTTGTGACCTTCTGTGGAAGATTTCATCTTGAACATGCAGATGCCCTCAGGTCCC 300
 QY 101 AspIleIleIleGlyValLeuGluPheSerSerIleHisTrpSerIleTrpIleThrVal 120
 DB 301 GACAAATCATAGAGAGTGTGAAATTCATCATCCATCCACCTCATATGATTAATCTGTA 360
 QY 121 ProLeuThrIleAspArgTYrIleAlaValCyseHisProLeuIleTYrHisIleThrValSer 140
 DB 361 CGTTAAACCATTTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTACACACGGTCTCA 420
 QY 141 TYrProAlaArgThrArgIleValIleValSerValTYrIleThrCysePheLeuThrSer 160
 DB 421 TACCAAGCCCGCACCCGAAAGTCAATGTAAGTTCATCACTCTCTCTGACACAGC 480
 QY 161 IleProTYrTYrTYrTrpProAsnIleTrpThrGluAspTYrIleSerThrSerValHis 180
 DB 481 ATCCCTCTTACTGTGTGGCCCAACATCTGAGCTGAAAGTCAATCAACACCTCTGTGCAT 540
 QY 181 HisValLeuIleTrpIleHisCysePheThrValTYrLeuValProCyseSerIlePhePhe 200
 DB 541 CACGCTCATCTGAGTACCATCGTCTTACCCGTACTGCTGTGCGCTCTCATCTTCTTC 600
 QY 201 IleuAsnSerIleIleValTYrIleValArgIleSerAsnPheArgIleuArgIly 220
 DB 601 ATCTGAATCATCAATCTGTGTACAGCTCAGAGAGAAAGCAATTTTCCTCCGTGGC 660
 QY 221 TYrSerThrGlyIleThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIleu 240
 DB 661 TACTCCAGGGGAAGAACACCGCCATCTTGTACCATTTACTTCCTTGGCACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTYrHisIleuTYrGlyAlaProIleGlnAsnArg 260
 DB 721 TGGGCCCGCCGATCATCATGATCTTTTACACCTCTTATGGGGCGCCCATCCAGAACCGC 780
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaIleuAsnThrAlaIle 280
 DB 781 TGGCTGTACATCATCTTCCGACATTTGCCAATCTGCAATCTGACCTTGTGAACACAGCATC 840
 QY 281 AsnPhePheLeuTYrCysePheIleSerIleArgPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACTCTCTCTACTGCTTACATCAGCAAGCGGTTCCGACCATGGCACCGCCACGCTC 900
 QY 301 IyeAlaPhePheLeuCyseGlnIleGlnProValGlnPheTYrThrAsnHisAsnPheSer 320
 DB 901 AAGGCTTTCTTCAAGGCAAGCAAACTGTACAGTTCTTACACCAATATATTAATCTTTC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCyseIleuMetLeuValTYr 340
 DB 961 ATTAACAAGTAGGCCCTCGATCTCGCCGGCAACTCACTGATCAAGATGCTGGGTGAC 1020
 QY 341 GlnTYrAspIleAsnGlyIleAspProIleIlys 350

Db 1021 CAGTATGACAAAATGSAACCTATATAA 1050

Search completed: December 14, 2004, 20:35:21
Job time : 421 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:57:16 ; Search time 81 Seconds
(without alignments)
3097.635 Million cell updates/sec

Title: US-09-995-225b-16

Sequence: 1 MHTTHAHLANSLSLWSPG.....CINMLVYQYDKNGKPIKVP 353

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO.spool/US09995225/Runat_14122004_110734_18093/app_query.fasta_1.519
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -INITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR_SCORE=EPC -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09995225 @CEN 1.1 54 @runat_14122004_110734_18093 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	241	12.9	1182	1	US-07-629-1041-2
2	241	12.9	1182	1	US-07-629-1041-1
3	238.5	12.7	1194	2	US-08-288-663A-1
4	238.5	12.7	1229	2	US-08-288-663A-2
5	238.5	12.7	1229	4	US-09-016-434-1428
6	232	12.4	1248	4	US-09-545-944-1
7	231.5	12.4	1495	4	US-09-016-434-1190
8	231.5	12.4	1495	4	US-09-023-655-1021
9	231.5	12.4	2156	4	US-08-012-988A-1
10	231.5	12.4	2156	4	US-09-023-655-1247
11	223	11.9	1600	4	US-09-029-027B-1
12	222.5	11.9	2772	4	US-09-919-039-171

13	220	11.8	2706	3	US-08-676-351-1	Sequence 1, App1
14	219.5	11.7	1317	4	US-09-016-434-1446	Sequence 1446, Ap
15	219	11.7	1452	1	US-08-149-093A-3	Sequence 3, App1
16	219	11.7	1452	1	US-08-911-245-3	Sequence 3, App1
17	219	11.7	1452	1	US-08-553-058C-3	Sequence 3, App1
18	219	11.7	1452	2	US-08-514-451A-3	Sequence 3, App1
19	219	11.7	1452	3	US-09-170-331-3	Sequence 3, App1
20	219	11.7	1452	3	US-09-510-473-3	Sequence 3, App1
21	219	11.7	1452	4	US-09-048-916B-3	Sequence 3, App1
22	218	11.6	1137	3	US-09-045-583-6	Sequence 6, App1
23	218	11.6	1137	3	US-09-534-185-6	Sequence 6, App1
24	218	11.6	1664	4	US-09-045-583-4	Sequence 4, App1
25	218	11.6	1664	4	US-09-534-185-4	Sequence 4, App1
26	217	11.6	1330	3	US-08-147-592A-5	Sequence 5, App1
27	217	11.6	1330	3	US-08-292-694A-5	Sequence 5, App1
28	217	11.6	1567	3	US-08-889-108-16	Sequence 16, App1
29	217	11.6	1567	5	PCT-US94-10358-16	Sequence 16, App1
30	217	11.6	2600	4	US-08-986-209A-1	Sequence 1, App1
31	217	11.6	2706	2	US-08-454-549-1	Sequence 1, App1
32	217	11.6	2706	3	US-08-454-552-1	Sequence 1, App1
33	216.5	11.6	1225	4	US-09-016-434-1235	Sequence 1235, Ap
34	214.5	11.5	1317	1	US-08-153-848-45	Sequence 45, App1
35	214.5	11.5	1317	3	US-09-299-843A-45	Sequence 45, App1
36	214.5	11.5	1317	3	US-09-088-337B-45	Sequence 45, App1
37	214.5	11.5	1317	5	PCT-US93-11151-45	Sequence 45, App1
38	213.5	11.4	1275	4	US-09-341-446B-7	Sequence 7, App1
39	213	11.4	1664	4	US-09-582-224A-5	Sequence 5, App1
40	213	11.4	1664	4	US-09-023-655-1213	Sequence 1213, Ap
41	213	11.4	1679	3	US-09-517-605-14	Sequence 14, App1
42	213	11.4	1737	1	US-08-202-056-4	Sequence 4, App1
43	213	11.4	1737	1	US-08-076-093A-3	Sequence 3, App1
44	213	11.4	1737	1	US-08-701-265-3	Sequence 3, App1
45	213	11.4	1737	2	US-08-284-586-3	Sequence 3, App1

ALIGNMENTS

RESULT 1
US-07-629-1041-2
Sequence 2, Application US/076291041
Patent No. 5286621
GENERAL INFORMATION:
APPLICANT: Gershengorn, Marvin C
TITLE OF INVENTION: PITUITARY TRH RECEPTOR
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/629,1041
FILING DATE: 19901214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: CRF D - 995
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 2 :
SEQUENCE CHARACTERISTICS:
LENGTH: 1182 base pairs
TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (cDNA)
US-07-629-1041-2

Alignment Scores:
Pred. No.: 9,88e-17      Length: 1182
Score: 241.00           Matches: 85
Percent Similarity: 41.78%      Conservative: 70
Best Local Similarity: 22.91%    Mismatches: 121
Query Match: 12.87%           Indels: 96
DB: 1                    Gaps: 15

US-09-995-225b-16 (1-353) x US-07-629-1041-2 (1-1182)

QY 19 ProGlySerAlaCysGlyLeuGlyPhe---ValProValValTyrTyrSerLeuLeuLeu 37
DB 46 CCACACAGCAGCTGTGGCCCTCGAGTACAGGTGGTTACCATCTTACTTGTGTGCTATTATT 105
QY 38 Cys---LeuGlyLeuProAlaAenIleuThrValIleIleuSerGlnLeuValAla 56
DB 106 TGTGGACTGGGCACTGTGGCCACATCATGTACTGCTGGTGTGATGACAGACAAAGCAC 165
QY 57 ArgArgGlnLysSerSerTyrAenTyrLeuLeuAlaLeuAlaAlaAspIleLeuVal 76
DB 166 ATGAGAACCCCTCAAACTGT---TACCTGTAAAGTCTGGCTGGCAGATCTCATGT 222
QY 77 LeuPhePheIleValPheValAAspPheLeuLeuGlnAspPheIleuSerMetGlnMet 96
DB 223 CTG-----GTCCTGCGAGGACTC 240
QY 97 ProGlnValProAspLysIle-----GTCCTGCTATGGCTATGGCTATGGCTGCTGCTG 103
DB 241 CCCACATATACCCAGATATCTATGGTCTGCTATGGCTATGGCTATGGCTGCTGCTGCTG 300
QY 104 IleGluValLeuGlnLysSerSerIleHisThrSerIleTyrIleThrValProLeuThr 123
DB 301 ATTACATATCTCCAGTACCTAGGCAATTAAATGATCTTCAATGTTCAATAACGCGCTTACC 360
QY 124 IleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSerTyrProAla 143
DB 361 ATGGAAGGATACATACATCTGTCTACCCCATCAAGCCAGTTTCTGCAACGTTTTC 420
QY 144 ArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSerIleProLys 163
DB 421 AAGGCCAAAAATCATCATCTTGTCTGGCC-----TTCACATTCATTACGT 471
QY 164 TyrTyrTyr-ProAenIleTyrThrGlnAspTyrIleSerThrSerVal---HisIleVal 182
DB 472 ATGCTCTGCTTCTCTGCTGCTG-----ATCTCAATCAGCAGCCTTACAAAAC 519
QY 182 IleuIleTyrIleHisCys-----PheThrValTyrIle 193
DB 520 GCTGTGTGTGTTTCC-TGTGGCTACAGATCTCCAGAACTACTACTACCATTTACT 578
QY 193 uValProCysSerIlePhePheIleLeuAsnSerIle-----IleValTyr----- 208
DB 579 AATGACACTTGGTGTCTTTATGTGTGCAATGATCTGGCCACTGTGCTTATGAT 638
QY 208 ----- 208
DB 639 TATAGTAGAATCTCTTCTTAACCCCATCTTCCAGACCCCTAAAGAAACTTAAGAT 698
QY 209 -----LysLeuArgLysSerAsnPheArgLeuArgLys----- 220
DB 699 GTGGAAAAATGATCATCATTCAGAACACAAAGATTGAAATTAATGACCAACAGATG 758
QY 221 -----TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrse 235
DB 759 CTTCAACAGCACTGATCTTCCAGAGACAGATCCCAAGATGCTCGCAGTGTGTAT 818
QY 235 rIlePheAlaThrLeuThrAlaPro---ArgIleIleMetIleLeuTyrHisLeuTyrG1 254
DB 254 ----- 254

```

```

DB 819 TCTGTTGGCCCTTTATGATGCCCTTACAGACTCTTGTGTGTGCACTATTCTCTC 878
QY 254 yAlaProIleGlnAsnArgTyrLeuValHisIleMetSerAspIleAlaAsnMetLeuAl 274
DB 879 CAGCCCTTTCCAGAAAAATTGGTCTGCTC-----TTTGGAGAAATTTGCAT 926
QY 274 AleuLeuAenThrAlaIleAsnPheLeuTyrCysPheIleSerLysArgPheArgTh 294
DB 927 TTATCTCAACAGATGCGATCAACCCAGTATTTACAACTCATGCTCAGAAATTTGCT-- 984
QY 294 rMetAlaAlaThrLeuLysAlaPhePheLysCysGlnLysGlnProValGlnPheTyr 314
DB 985 -----GCAGCCTTCAGAGACCTGTGCATTTGCAAGAGAACCCACAGAAAAAGC 1034
QY 314 rThrAsnHisAenPheSerIleThrSerSer 324
DB 1035 TGCTAAGTACAGTGTGGCCCTTAATTAACGT 1065

```

```

RESULT 2
US-07-629-1041-1
; Sequence 1, Application US/076291041
; Patent No. 5288621
; GENERAL INFORMATION:
; APPLICANT: Gershengorn, Marvin C
; APPLICANT: Straub, Richard E
; TITLE OF INVENTION: PITUITARY TRH RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/629,1041
; FILING DATE: 19901214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D - 995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 268-1951
; TELEFAX: (203) 268-1951
; INFORMATION FOR SEQ ID NO: 1 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1752 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (cDNA)
US-07-629-1041-1

```

```

Alignment Scores:
Pred. No.: 1,79e-16      Length: 1752
Score: 241.00           Matches: 85
Percent Similarity: 41.78%      Conservative: 70
Best Local Similarity: 22.91%    Mismatches: 121
Query Match: 12.87%           Indels: 96
DB: 1                    Gaps: 15

```

```

US-09-995-225b-16 (1-353) x US-07-629-1041-1 (1-1752)

QY 19 ProGlySerAlaCysGlyLeuGlyPhe---ValProValValTyrTyrSerLeuLeuLeu 37
DB 304 CCACACAGCAGCTGTGGCCCTCGAGTACAGGTGGTTACCATCTTACTTGTGTGCTATTATT 363

```

```

QY 38 Cys---LeuGlyLeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAla 56
DB 364 TGTGACCTGGGCAATGTGGCAACATCATGTAGTCTGTGGTGCATGAGAACAAAGCAC 423
QY 57 ArgArgGlnIleSerSerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuVal 76
DB 424 ATGAGAACCCCTCAAAACTGT---TACCTGTAGAGTGTGGCTGTGGAGAGTCTCATGGTT 480
QY 77 LeuPhePheIleValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMet 96
DB 481 CTG-----GTCGCTGCAGAGACTC 498
QY 97 ProGlnValProAspIleVal 103
DB 499 CCCAACATTAACGACATCATGTATGTTCTGTGGTCTATGAGCTATGTTGGCTGCTTGC 558
QY 104 IleGlnValLeuGlnPheSerSerIleIleThrSerIleThrIleThrValProLeuThr 123
DB 559 ATTACATATCTCCAGTACCTAGGCAATTATGATCTTTCATGTTCAATTAAGGCTTTTACC 618
QY 124 IleAspArgTyrIleAlaValCysHisProLeuIleYrHisThrValSerTyrProAla 143
DB 619 ATGGAAGGTACATAGCAATCTGTCAACCCCATCAAGCCAGATTCTCTCAGGTTTTC 678
QY 144 ArgThrArgIleValIleValSerValTyrIleThrCysPheLeuThrSerIleProTyr 163
DB 679 AGAGCCCAAAAATCATCATCTTGTCTGGGCC-----TTCACATCCATTACTGT 729
QY 164 TyrTrpTyr-ProAsnIleThrPthrGlnAspTyrIleSerThrSerVal---HisHisVal 182
DB 730 ATGCTGTGTTCTTCCGCTGC-----ATCTCAACATCAGCAGCTACCAAAAC 777
QY 182 IleuIleThrIleHisCys-----PheThrValTyrIle 193
DB 778 GCTGTGTGGTTTC--TGTGGCTAACAGATCTCCAGAACTACTACTACCTATTATTAAGT 836
QY 193 ValAlProCysSerIlePhePheIleLeuAsnSerIle-----IleValTyr----- 208
DB 837 AATGACCTTGTGGTCTTATGTATGTGTGCCAATGATCTGGCCACGTGTCTTATGAGATT 896
QY 208 ----- 208
DB 897 TATAGCTAGAAATCCTCTTAAACCCCATCTTCAGACCTTAAGAAAACCTTAAGAT 956
QY 209 -----LysLeuArgArgIleSerAsnPheArgLeuArgIle----- 220
DB 957 GTGAAAAAATGACTCATCATCATCAGAAACAAGATTGAATTAAATGCCAACAAAGATG 1016
QY 221 -----TyrSerThrGlySerThrAlaIleLeuPheThrIleThrIle 235
DB 1017 CTTCAACAGACGTATCTTCTTCAAGAAAGCTCAGCAAGATGCTCGAGTGTGTAAT 1076
QY 235 rIlePheAlaThrLeuThrAlaPro---ArgIleIleMetIleLeuTyrHisLeuTyrGln 254
DB 1077 TCGTTTGGCTTTCATTCATGAGTCCCTTACAGAGACTCTGTGGTGTCAATCTTCTCTC 1136
QY 254 ValAlaProIleGlnAsnArgTyrLeuValHisIleMetSerAspIleAlaAsnMetLeuAl 274
DB 1137 CAGCCCTTTCAGAAATAATGTTCTTCTGCTC-----TTTTCAGAAATTTCAT 1184
QY 274 AleuLeuAsnThrAlaIleAsnPhePheLeuTyrCysPheIleSerIleYrAspPheArgThr 294
DB 1185 TTATCTCAACAGTCCATCAACCCAGATTTTCAACCTCATGTCTCAAGAAATTTCTGT- 1242
QY 294 rMetAlaAlaAlaThrLeuIleValAspPhePheIleCysGlnIleGlnProValGlnPheTyr 314
DB 1243 -----GAGAGCTTCAGGAAGCTTCGCAATTCGACAGACAGAACCCACAGAAAAGC 1292
QY 314 rThrAsnHisAsnPheSerIleThrSerSer 324
DB 1293 TGTAACTACAGTGTGGCCCTTAATTAATTAAGT 1323

```

```

US-08-288-663A-2
/ Sequence 2, Application US/08288663A
/ Patent No. 587986
/ GENERAL INFORMATION:
/ APPLICANT: HINDUMA, Shuji
/ APPLICANT: HOSOGA, Masaki
/ APPLICANT: ONDA, Haruo
/ TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
/ TITLE OF INVENTION: AND USE
/ NUMBER OF SEQUENCES: 15
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
/ STREET: 130 Water Street
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/288,663A
/ FILING DATE: 09-AUG-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 198309/1993
/ FILING DATE: 10-AUG-1993
/ APPLICATION NUMBER: 286986/1993
/ FILING DATE: 16-NOV-1993
/ APPLICATION NUMBER: 325215/1993
/ FILING DATE: 22-DEC-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Resnick, David S
/ REGISTRATION NUMBER: 34,235
/ REFERENCE/DOCKET NUMBER: 44612
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1194 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-288-663A-2
Alignment Scores:
Pred. No.: 1,88e-16 Length: 1194
Score: 238.50 Matches: 85
Percent Similarity: 41.21% Conservative: 72
Best Local Similarity: 22.31% Mismatches: 130
Query Match: 12.74% Indels: 95
DB: 2 Gaps: 13
US-09-995-225b-16 (1-353) x US-08-288-663A-2 (1-1194)
QY 2 GlnHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTyrTrpSerProGlySer 21
DB 32 AAACACAGCTTCAGCCAC-----GAGCAGTGTGGCTTGAATACC 73
QY 22 AlaCysGlyLeuGlyIlePheValProValIleTyrTyrSerLeuLeuLeuCysLeuGlyLeu 41
DB 74 AGG-----TGGTCACCATCTTACTTACTTATTTGTGGC-CTGGGCATT 120
QY 42 ProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnIleSer 61
DB 121 GTAGGCAACATCATGTAGTCTGTGTGCATGAGAACCAAGACATGAGAGCCCAAC 180
QY 62 SerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValIlePhePheIleVal 81

```



```

Db      328 TATTGGGAATTAATGATCCTTTGTCATAAAGACCTTTACATTGAGAGTACATA 387
Qy      129 AlAlAlCySHSPoleuLyTyRHISThValSerTyProAlaArgThArgLyVal 148
      388 GCATCTGTGACCCCAAGGCCAGTCTTCTGCACATTTTCCAGAGCCAAAGAT 447
Qy      149 lIeValSerValTyRlIeThCysPheLeuThSerIleProTyTyRTPPro----- 166
      448 ATCATCTTCTGCGGT-----TTCAATCTTCTTACTGATGCTCTGCTTTC 498
Db      167 -----ProAnIleTyRThr--GluAspTyRlIeSerThSerValHISVal 182
Qy      499 TTGCTGATCTCATATATTACCAAGAGTCTATCTGATTCCTTGCTGCTCAAG 558
Db      183 LeuIleTyRlIeHIScysPheThValTyRLeuValProCysSerIlePhePheIleu 202
      559 ATCTCCAGGATTACTACTACCTATTACTTAATGACTTTGGTGTCTTATGTGTG 618
Qy      203 AsnSerIleIleValTyRlySLeu----- 210
      619 CCAATGATCTGGCTACCGTCTCTATGATTCATAGTACTAGATCCTTTTCTTAATCCC 678
Qy      211 ----- 213
      679 ATTCCTTCAGATCTTAAGAAAGAACTTAAGACATGAGAAATGATTCAACCCATCAGAAC 738
Db      214 SerAsn-----PheArgLeuArgLyTySerThArgLy 224
      739 ACAAACTGAAATGAAATACCTTAATAGATGTTTCAACGACAGATCTTCAAGAAAG 798
Qy      225 LysThThAlaIleuPheThRlIeThSerIlePheAlaThRleuTyRlaPro--- 243
      799 CAGGTACCAAGATCTGGCAGGTGTTGATTCCTTTGCCCTTTATGATGCCCTAC 858
Db      244 ArgIleIleMetIleuTyRHISLeuTyRGlYAlaProIleGlnAsnArgTyRleuVal 263
      859 AGACTCTGTGTGTTCAACTCATTTCTCTCCAGCTTTCACAAATAATGTGTTTGG 918
Qy      264 HsIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThAlaIleAsnPhe 283
      919 CTC-----TTTGCAAGATTTTGCAATTTATCTGACAGTCCCAACCCGGTG 966
Db      284 LeuTyTyCysPheIleSerIleArgPheArgThMetAlaAlaThRleuTyRlaPhe 303
      967 ATTACAAATCTCATGTCCCAAGAAATTCGCT-----GCAGCTTCAGAAAGCTC 1014
Qy      304 PheTyCysGlnIleGlnProValGlnPheTyThRAsnHISAsnPheSerIleThSer 323
      1015 TGCACCTGCAAGCAAGAGCCAAAGAAACCTGCTAACTACAGTGTGCGCTTAATTA 1074
Qy      324 Ser 324
      1075 AGC 1077
Db

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1248
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-545-944-1

Alignment Scores:
Score: 1,04e-15 Length: 1248
Percent Similarity: 232.00 Matches: 79
Best Local Similarity: 44.32% Conservative: 77
Query Match: 22.44% Mismatches: 112
DB: 12.39% Indels: 84
Gaps: 15

US-09-995-225b-16 (1-353) x US-09-545-944-1 (1-1248)

Qy      27 PheValProVal---ValTyTySerIleuLeuCySLeuGlyLeuProAlaAsnIle 45
      133 TTCCTCCCGGTCTGTGTGTATGTGCAATTTTGTGTGGGGGTGATGGCAATGTC 192
Db      46 LeuThValIleIleLeuSerGlnLeuValAlaArgArgLysSerSerTyRAsnTyR 65
      193 CTGCTGCTGCTGTATCTGACAGCAGCAGCTATGAAGAGCCCACTAC---TAC 249
Qy      66 LeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIleValPheValAspPhe 85
      250 CTCCTTACCTGGCGCTCTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 297
Qy      86 LeuLeuGluAspPheIleLeuAsnMetGlnMetPro----- 97
      298 CCCCTGAGAGTCTATGATGTGGCACTACCTTTCTTGTGGGCCCGGTGGCTGC 357
Db      98 GlnValProAspLysIleIleGlnValLeuGlnPheSerSerIleHISThSerIleTyR 117
      358 TACTTCAAGAGCGCCCTCTTGAAGCGGTGCTTCCGCTCAGC-----CTCAGC 408
Qy      118 lIeThRValProLeuThRlIeAspArgTyRlIeAlaValCySHSPoleuLyTyRHIS 137
      409 ATCAACACC---GTGAGGTGAGCGCTTACCTGCATCTCAACCCGTTCCGCCAAG 465
Qy      138 ThValSerTyProAlaArgThArgLyValIleValSerValTyRlIeThCysPhe 157
      466 CTGAGAGACCCCGCGCGCGCCCTCAGATCTCGGATCGTGTGGGCTTCCGCTG 525
Db      158 LeuThSerIlePro-----TyTyTyRTPProAsn--- 168
      526 CTCCTTCCCTGCCCAACACAGCATGCATGCATGAATTCACATCTTCCCAATGGG 585
Qy      169 -----lIeTyRThArgLyTyR 174
      586 TCCCTGTCAGGTTGGCCACCTGACGTCATCAAGCCCATGTG----- 633
Db      175 lIeSerThSerValHISValLeuIleTyRlIeHIScysPheThValTyRleuVal 194
      634 -----ATCTCAATTTTCATCATCCAGGTACCTCTTCTTACTTACTCTCTC 681
Qy      195 ProCysSerIlePhePheIleuAsnSerIleIleValTyRlySLeuArgArg----- 212
      682 CCCATGACTGTATCACTGTCTCTACTACATGCTGACACTGAAGAAAGAACAA 741
Qy      213 -----LysSerAsnPheArgLeuArgLyTyRserThArgLyTyRThR 227
      742 TCTCTGAGCAATGAAGGAATGCAATATTCAGAACCTGCGCAAAATAGTCAAC 801
Db      228 AlAlIleuPheThRlIeThSerIlePheAlaThRleuTyRlaProArgIleIleMet 247
      802 AAGATGCTGTTGTGTGCTTGAATGTTGCTATGTTGGGCCCGCTTCCACATTTGAC 861
Qy      248 lIeLeuTyRHISLeuTyRGlYAlaProIleGlnAsnArgTyR----- 261
      862 CCACTCTTCTTACCTTT-----GTGAGAGAGTGAATGATCCCTGCTGCT 909

```

```

Oy 262 -----LeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThr 278
Db 910 GTGTTCAACCTCGTTCAGTGTGTGTCAGGTGC-----TATTCTTACTTGAGCTCA 960
Oy 279 AlaIleAsnPhenPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAla 298
Db 961 GCTGTCAACCCCATATATCTATTACTACTGCTGTGCGCGCTTCACGACGATTCAGAAAT 1020
Oy 299 ThrLeuValAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsn 318
Db 1021 GTGATCTCTCTCTTCTTCCACAAA-----CAGTGCACTCCAGCATGAC 1062
Oy 319 PheSerIleThrSerSerProTryptIleSerProAla 330
Db 1063 -----CCACACATGGCCACCTGCC 1080

RESULT 7
US-09-016-434-1190
; Sequence 1190, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1190:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENEBANK
; CLONE: g179984
; US-09-016-434-1190

Alignment Scores:
Pred. No.: 1.56e-15 Length: 1495
Score: 231.50 Matches: 81
Percent Similarity: 44.81% Conservative: 57
Best Local Similarity: 26.30% Mismatches: 105
Query Match: 12.37% Indels: 65
DB: 4 Gaps: 13

```

```

QY      22  AlaCysGlyLeucllypheaIProValValTyrTyrSerLeuLeuLeuCysLeuGlyLeu  41
Db      91  GCCCTTGGGGCCCACTGGCTCCCTCCCTG--TACCTCTTGTAATTTGTTCATTGGCTG  147
QY      42  ProAlaasnIleuThrValIleIleuSerGlnLeuValAlaArgGlnLysSer  61
Db      148  GTTGGAAACATCTCTGGTGGTCTCGGTCCTGTGGCAATACAGAGGCTAAACAAATGAC  207
QY      62  SerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPheIleVal  81
Db      208  AGCATC---TACCTCTGAAACCTGGCACTTCTGACCTCTTCCTGTTACAGCTTCCC  264
QY      82  Phe---ValAspPhe---LeuGlnAspPheIleLeuAsnMetGlnMetProGlnVal  99
Db      265  TTTCGATGCATCAACAGTTGAAGATGCTGGGTTTTTGGTGAATGCATG-----  315
QY      100  ProAspLysIleIleGlnValLeuGlnPheSerSerIleIleIleSerIleTyrIleThr  119
Db      316  ---TGTAAGATCTCTCTGGTATTATTAATACAGAGCTGTACAGAGATCTTTTCATC  372
QY      120  ValProLeuThrIleAspArgTyrIleAlaValCysHis-----ProLeuGlyTyr  156
Db      373  ATCTGTGTGACGATTACAGTAACCTGGCACTGGCACTGCACGCGGTGTTGCCCTTGGGCA  432
QY      137  HisThrValSerTyrProAlaArgThrArgLysValIleValSerValTyrIleThrCys  156
Db      433  CGAGACGCTCACTTTGGTGTCTATCAACAGCATCATCATTTGGGCGCTGGCATC-----  486
QY      157  PheLeuThrSerIlePro---TyrTyrTyrTyrProAsnIleTyrThrGlnAspTyrIle  175
Db      487  ---TTGGCTTCAGATGCCAGGCTTATACTTTTCCAGAACCCATGGGAATTCATCATCAC  543
QY      176  SerThrSerValHisHis-----  180
Db      544  ACCTGACGCTTCATCTTCTCTGACGAAAGCTTACGAGAGTGAAGCTGTTCAAGCTCTG  603
QY      182  -----ValleuIleTyrIleHisCysPheThrVal  191
Db      604  AAACGTAACTCTTTGGGCTGGATTTGCTTTGTGTGCATGATCATCTGCTATACAC---  666
QY      192  TyrLeuValProCysSerIlePhePheIleLeuAsnSerIleIleValTyrLysLeuArg  211
Db      661  -----GGATTATTAAGAATTCTGCTAAGA  684
QY      212  ArgLysSerAsnPheArgLeuArgGlyLysSerThrArgLysSerThrAlaIleLeuPhe  233
Db      685  CGACCAAAATAG-----AAGAAATCCAAAGCTGCTCCGTTATATTTT  726
QY      232  ThrIleThrSerIlePheAlaThrLeuTyrAlaProArgIleIleMetIleLeuTyrHis  251
Db      727  GTCATCATATCATCTTTTCTCTTTTGGACCCCTTCAATTTGACATATCTTATTTCT  786
QY      252  LeuTyrGlyValAlaProIle-----GlnAsnArgTyrPheValHisIle  265
Db      787  GTTTTCCAGACCTCTCTGTTCACCATGATGTTGACAGACAGACATTTG---GACCTG  843
QY      266  MetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIleAsnPhePheLeuTyr  285
Db      844  GCTGTGCAMGTGACGAGGATGATCGCTTACACGCACTGCTGTGTCAACCCAGTATCTAC  903
QY      286  CysPheIleSerLysArgPheArg  293
Db      904  GCCTTGCTGTGTGAGAGGTTCCGG  927

RESULT 8
US-09-023-655-1021
; Sequence 1021, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
;   APPLICANT: Cocks, Benjamin G.
;   APPLICANT: Susan G. Stuart
;   APPLICANT: Jeffrey J. Seilhammer

```

```

/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
/ TITLE OF INVENTION: EXPRESSION
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ FILING DATE: US/09/023,655
/ CLASSIFICATION:
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ INFORMATION FOR SEQ ID NO: 1021:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1495 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g179984
/ US-09-023-655-1021

Alignment Scores:
Pred. No.: 1,55e-15 Length: 1495
Score: 231.50 Matches: 81
Percent Similarity: 44.81% Conservative: 57
Best Local Similarity: 26.30% Mismatches: 105
Query Match: 12.37% Indels: 65
DB: Gaps: 13

US-09-995-225b-16 (1-353) x US-09-023-655-1021 (1-1495)
QY 22 AlaCgGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGlyLeu 41
DB 91 GCCTTGGGGCCCACTGCTGCCCTCTG---TACTCTTGATATTGTCATTGGCTG 147
QY 42 ProAlaAsnIleLeuThyValIleIleLeuSerGlnLeuValAlaArgArgGlnLysSer 61
DB 148 GTTGAACATCTCTGCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 207
QY 62 SerTyrAsnTyrLeuLeuAlaLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 81
DB 208 AGCATC---TACTCTGTAACCTGCGCATTTTGTGACCTGCTCTTCTTTCACGCTTCC 264
QY 82 Phe---ValAspPhe---LeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnVal 99
DB 265 TTCTGATCGACTACAGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 315
QY 100 ProAspLysIleIleGlnValLeuGlnLubHeserIleHisThrSerIleTyrIleThr 119
DB 316 ---TGTAGATTCCTCTGCTGCTTATTATACACAGGCTTGTACACGAGATCTTTTGCATC 372
QY 120 ValProLeuThrIleAspArgTyrIleAlaValCysHis-----ProLeuLysTyr 136
DB 136

```

```

DB 373 ATCTGCTGACGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATG 432
QY 137 HisThrValSerTyrProAlaArgThrArgLysValIleValSerValTyrIleThrCys 156
DB 433 CGGACCGTACCTTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 486
QY 157 PheLeuThrSerIlePro---TyrTyrTyrTyrProAlaIleIleThrGlnAspTyrIle 175
DB 487 ---TTGCTTCCATGACGCTTATCTTATCTTCCAGACCCATGGAATTACTACACAC 543
QY 176 SerThrSerValHisIle-----ValIleIleTyrIleHisCysPheThrVal 181
DB 544 ACTCGACGCTTCACTTCTTCCACGAAAGCTTACGAGATGAGAGCTGTTTCAGGCTCG 603
QY 182 -----ValIleIleTyrIleHisCysPheThrVal 191
DB 604 AAACGACCTTGGGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 660
QY 192 TyrLeuValProCysSerIlePhePheIleLeuAsnSerIleIleValTyrLysLeuArg 211
DB 661 -----GGGATTATTAAGATTTGCTAAGA 684
QY 212 ArgLysSerAsnPheArgLeuArgGlyTyrSerThrGlyLysThrAlaIleLeuPhe 231
DB 685 CGACCAATGAG-----AAGAAATCCAAAGCTGTCCTGTTGATTTT 726
QY 232 ThrIleThrSerIlePheAlaThrLeuThrAlaProArgIleIleMetIleLeuTyrHis 251
DB 727 GTCATCATGATCATCTTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 786
QY 252 LeuTyrGlyAlaProIle-----GlnAsnArgTyrLeuValHisIle 265
DB 787 GTTTCCAAGACTTCCTGCTTCCACCATGATGATGATGATGATGATGATGATGATGATG 843
QY 266 MetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhePheLeuTyr 285
DB 844 GCTGTGCAAGTACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 903
QY 286 CysPheIleSerLysArgPheArg 293
DB 904 GCCTTCGTTGGTGAGAGTTCCGG 927

RESULT 9
US-08-012-988A-1
/ Sequence 1, Application US/08012988A
/ Patent No. 5652133
/ GENERAL INFORMATION:
/ APPLICANT: Murphy, Philip M.
/ TITLE OF INVENTION: Cloning and Expression of Human
/ TITLE OF INVENTION: Macrophage Inflammatory Protein-1 alpha (MIP-1)
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend Kourie and Crew
/ STREET: One Market Plaza, Stewart Tower, Suite 2000
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94610
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/012,988A
/ FILING DATE: 19930128
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Weber, Kenneth A.
/ REGISTRATION NUMBER: 31,677
/ REFERENCE/DOCKET NUMBER: 15280-118
/ TELECOMMUNICATION INFORMATION:

```


Percent Similarity: 44.81%
 Best Local Similarity: 26.30%
 Query Match: 12.37%
 Gaps: 4
 Conservative: 57
 Mismatches: 105
 Indels: 65
 Gaps: 13

US-09-995-225B-16 (1-353) x US-09-023-655-1247 (1-2156)

```

QY 22 AlaGgLyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGlyLeu 41
Db 153 GCGTTTGGGGCCCACTGCTGCCCCCTCTG--TACTCTCTGATTTGTCATGGCCCTG 209
QY 42 ProAlaAniLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLysSer 61
Db 210 GTTGAAGAACATCCGTGGTGTCTGCTGCTTGTGAATACAGAGGCTAAAGACATGACC 269
QY 62 SerTyrAniLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPheIleVal 81
Db 270 AGCATC--TACTCTGTAACCTGGCCATTTCTGACCTGCTTCTGTCAGCTTCCC 326
QY 82 Phe--ValAspPhe--LeuLeuGluAspPheIleLeuAniMetGlnMetProGlnVal 99
Db 327 TTTGTGATCGACATCAAGATTGAAGATGACTGGGTTTGTGTATGCCATG-- 377
QY 100 ProAspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTyrIleThr 119
Db 378 --TGTAAATCCTCTCTGGGTTTATTATACAGAGCTTGTACAGCGAGATCTTTTCATC 434
QY 120 ValProLeuThrIleAspArgTyrIleAlaValCysHis-----ProLeuLysTyr 136
Db 435 ATCTGTGTACGATTGACAGATTACCTGGCCATCTGTCACCGCGTGTGGCTTGGCGGCA 494
QY 137 HisThrValSerTyrProAlaArgThrArgValIleValSerValTyrIleThrCys 156
Db 495 CGAGCCGTCACCTTTGGTGTGCATCACGACATCATATTGGCCCTGGCCATC----- 548
QY 157 PheLeuThrSerIlePro--TyrTyrTyrTyrProAniLeuTyrThrGluAspTyrIle 175
Db 549 --TTGGCTTCACAGCCAGGCTTATCTTTCACAGACCCAAATGGAAATTCCTACACAC 605
QY 176 SerThrSerValHisHis-----ValLeuIleTyrIleHisCysPheThrVal 181
Db 606 ACCTGAGAGCTTACCTTCTCCTACGAAACCTACGAGAGTGGAGCTGTTTCAAGGCTCTG 665
QY 182 -----ValLeuIleTyrIleHisCysPheThrVal 191
Db 666 AAATGAACTCTTGGGCTGTATTCCTTGTGTGATGATCATCTGTACACA--- 722
QY 192 TyrLeuValProCysSerIlePhePheIleLeuAniSerIleIleValTyrLysLeuArg 211
Db 723 -----GGGATTATTAAGATTCTGTCTAAGA 746
QY 212 ArgLysSerAsnPheArgLeuArgLysTyrSerThrGlyLysThrThrAlaLeuPhe 231
Db 747 CCACCAAAATGAG-----AAGAAATCCAAAGCGTCCGTTGATTTT 788
QY 232 ThrIleThrSerIlePheAlaThrLeuThrPalaProArgIleIleMetIleLeuTyrHis 251
Db 789 GTCAATCATGATCATCTTTTCTCTTGTGACCCCTCAAAATTTCAATTAATTTCT 848
QY 232 LeuTyrGlyAlaProIle-----GlnAniArgTyrPheValHisIle 265
Db 849 GTTTTCCAAAGACTTCTGTTCACCATGATGTGACAGAGACATTTG--GACCTG 905
QY 266 MetSerAspIleAlaAsnMetLeuAlaLeuLeuAniThrAlaIleAsnPhePheLeuTyr 285
Db 906 GCTGTGCAAGTACGAGAGTATCGCTTACACGACACTGCTGTGTCAACCCAGTATCTAC 965
QY 286 CysPheIleSerLysArgPheArg 293
Db 966 GCGTTCTGTGTGAGAGGTTCCGG 989

```

RESULT 11
 US-09-029-027B-1

```

/ Sequence 1, Application US/09029027B
/ Patent No. 6441133
/ GENERAL INFORMATION:
/ APPLICANT: Walker, Philippe
/ TITLE OF INVENTION: No. 6441133el TRH Receptor
/ FILE REFERENCE: Walker appl.
/ CURRENT APPLICATION NUMBER: US/09/029,027B
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 1600
/ TYPE: DNA
/ ORGANISM: Rat
US-09-029-027B-1

Alignment Scores:
Pred. No.: 1,466-14 Length: 1600
Score: 223.00 Matches: 79
Percent Similarity: 43.12% Conservative: 62
Best Local Similarity: 24.16% Mismatches: 97
Query Match: 11.91% Indels: 90
Gaps: 13

US-09-995-225B-16 (1-353) x US-09-029-027B-1 (1-1600)
QY 31 ValTyrTyrSerLeuLeuLeuCys--LeuGlyLeuProAlaAniLeuThrValIle 49
Db 437 GCTTCTCAAGTCTCTGTGTGACACCTGGGCATTCGGGCATGCGATGATTTCTG 496
QY 50 IleLeuSerGlnLeuValAlaArgArgGlnLysSerSerTyrAn--TyrLeuAla 68
Db 497 GTGGG-----CTGACCTCAAGTACATGACACACCCACCACTGCTACTGTCAGC 550
QY 69 LeuAlaAlaAlaAspIleLeuValLeuPhePheIleValPheValAspPheLeuGlu 88
Db 551 CTGGCCCTCGCTGACCTCTCTGCTG----- 577
QY 89 AspPheIleLeuAniMetGlnMetProGlnValProAspLysIle----- 103
Db 578 -----CTGGCTGGGGGTCTGCCCAATGTCTGACAGCCTAAGTGGGGCACTGGATC 628
QY 104 -----IleGluValLeuLysSerSerIleHisThrSer 115
Db 629 TATGAGAGTGTGCTGCTCTGGGCATCACTTCCAGTACCTGGGCATCAATGTCTCC 688
QY 116 IleTyrIleThrValProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLys 135
Db 689 TCCTTCTCAATCCCTGGCTTCACTGTGAGAGGATATATACCATTTGCCACCCATGAGA 748
QY 136 TyrHisThrValSerTyrProAlaArgThrArgLysValIleValSerValTyrIleThr 155
Db 749 GCACAGACCGTGTGACACTGTGGCCGGCCAAACGATCATGTGCAGGACATCTGGGGGCTC 808
QY 156 CysPheLeuThrSerIleProTyrTyrTyrTyrTyrProAniLeuTyrThrGluAspTyrIle 175
Db 809 ACGTCCCTCATATGCTCACTGCTGTCTCTCTG-----TGATCTCAATGTCCG 859
QY 175 eSerThrSerValHisHisValLeuIleTyrIleHis-----CysPheThrValTyr 192
Db 860 GACAAACGAGCCCTTGAATGTGCTTACAA--GTGCCCGAGAGACTTACTCTGCCCATTA 918
QY 192 rLeuValProCysSerIlePhePheIle-----LeuAniSerIleIleValTyrLys 209
Db 919 CTGTGCTGACTTGTGCTGTCTTTTCAATGACACCTTGTGTGACCTGTGTCTATAG 978
QY 209 sLeu----- 210
Db 979 GCTCATCGGAGAGATTATTATTCAGAGCCCGTTGTCCAGAACTGTGCAGAGAGAG 1038
QY 211 -----ArgArgLysSerAsn 216
Db 1039 GCAGCCCATGGGCAAGAGAGGCTGACACAGCAACTGCTCAGGGCCAAAGAGC----- 1093

```

QY 216 eaRgLeuAArgGlyTyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSer11 236
DB 1094 -----TCCAGGAAGCAGCGCCAGCATGCTGGCGTGTGTGTGCT 1137
QY 236 ePheAlaThrLeuThrAlaPro---ArgIleIleMetIleLeuTyrHisLeuTyrGlyAl 255
DB 1138 TTTTGCCGCTGCTGTGACCCCTTACCGCACATGCTGCTCAACTCTCTTGTGGCCCA 1197
QY 255 aProIleGlnAAsnArgTyrLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLe 275
DB 1198 GCGTTTCCGTGGACCCCTGGCTGCTGCG-----TTCTGCCGACCTGTGTCTA 1245
QY 275 uLeuAAsnThrAlaIleAsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMe 295
DB 1246 CACCAACAGCGCTGTCAACCCCTGCTGCTACAGCTGATGTCAAGAGTTCCGG----- 1300
QY 295 ValAlaAlaThrLeuLys 301
DB 1301 ---GCGGCTTCTCTGAAA 1315
RESULT 12
US-09-919-039-171
Sequence 171, Application US/09919039
Patent No. 6727066
GENERAL INFORMATION:
APPLICANT: Kaeer, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
FILE REFERENCE: PA-0035 US
CURRENT APPLICATION NUMBER: US/09/919, 039
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 60/222,113
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 401
SOFTWARE: PERL Program
SEQ ID NO 171
LENGTH: 2772
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6727066 331306.1
NAME/KEY: unsure
LOCATION: 2249, 2262
OTHER INFORMATION: a, t, c, g, or other
US-09-919-039-171
Alignment Scores:
Pred. No.: 3,8e-14 Length: 2772
Score: 222.50 Matches: 81
Percent Similarity: 43.91% Conservative: 56
Best Local Similarity: 25.96% Mismatches: 103
Query Match: 11.89% Indels: 72
DB: 4 Gaps: 13
US-09-995-225B-16 (1-353) x US-09-919-039-171 (1-2772)
QY 22 AlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCyLeuGlyLeu 41
DB 258 GCGTTTGGGGCCCACTGCTGCCCTCTG---TACTCTTGGTATTTGTCATTTGGCTG 314
QY 42 ProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLysSer 61
DB 315 GTTGAACAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 374
QY 62 SerTyrAsnTyrLeuLeuAlaLeuAlaAlaAlaAspIleLeuValLeuPhePheIleVal 81
DB 375 AGCATC---TACCTCTGAACCTGGCCATTCTGACCTGCTTCTCTTCCAGGCTTCCC 431
QY 82 Phe---ValAspPhe---LeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnVal 99
DB 432 TTCTGATGACACTACAAAGTTGAAGATGATGCGGTTTGTGTATGCAAG----- 482

QY 100 ProAspLysIleIleGlnValLeuGluPheSerSerIleHisThrSerIleTyrIleThr 119
DB 483 ---TGTAAATCTCTCTGGGTTTTTATTATACAGCGCTGTATACAGCAATCTTTTCATC 539
QY 120 ValProLeuThrIleAspArgTyrIleAlaValCysHis-----ProLeuLysTyr 136
DB 540 ATCTGCTACAGATTACAGTACCTGGCCATGCTCCAGCGCTGTGGTGGCTGGCGCA 599
QY 137 HisThrValSerTyrProAlaArgThrArgLysValIleValSerValTyrIleThrCys 156
DB 600 CGAGCCGTACTTTTGGTGTATCATCACAGATCATATATTGGGCCCTGGCGCATC----- 653
QY 157 PheLeuThrSerIlePro---TyrTyrTyrTyrProAsnIleTyrThrGluLysTyrIle 175
DB 654 ---TTGGCTTCATGTCAGAGCTTATCTTTCCTCCAGACCAATGGGAATTCATCTCACAC 710
QY 176 SerThrSerValHisHis----- 181
DB 711 ACCTGAGCCTTCACTTCTCTCTACAGAACCTTACAGAGCTGAGCTTTTCAAGCTCTG 770
QY 182 -----ValLeuIleTyrIleHisCysPheThrVal 191
DB 771 AAAGTAACTCTTGGGCTGTGATGCTTGTGTGTATGATCATCTGCTACACA--- 827
QY 192 TyrLeuValProCysSerIlePhePheIleLeuAsnSerIleIleValTyrLysLeuArg 211
DB 828 -----GGGATTATAAAGATTCTGTCAAG 851
QY 212 ArgLysSerAsnPheArgLeuArgLysTyrSerThrGlyLysThrAlaIleLeuPhe 231
DB 852 CGACCAATGAG-----AAGAAATCCAAAGCTCTCCGTTGATTTT 893
QY 232 ThrIleThrSerIlePheAlaThrLeuThrAlaProArgIleIleMetIleLeuTyrHis 251
DB 894 GTCATCATGATCATCTTTTCTCTTGTGACCCCTACAAATTGACTATTAATTCTT 953
QY 252 LeuTyrGlyAlaProIleGlnAsnArgTyrLeuValHisIleMetSer----- 267
DB 954 GTTTTC-----CAGACTTCTCTGTTCACCCATGAGTGTAGAGAGACAGACAT 1001
QY 268 -----AspIleAlaAsnMetLeuAlaLeuAsnThrAlaIleAsnPh 282
DB 1002 TTGACCTGGGCTGTGCAAGTACGAGGTGATGCTTACAGCAGCTGCTGTCAACC 1061
QY 282 ePheLeuTyrCysPheIleSerLysArgPheArg 293
DB 1062 AGTATCTACGCTTGTGTTGTGAGAGGTTCCGG 1095
RESULT 13
US-08-676-351-1
Sequence 1, Application US/08676351C
Patent No. 6046026
GENERAL INFORMATION:
APPLICANT: EPIPLER, CECIL
APPLICANT: OZENERGER, BRADLEY
TITLE OF INVENTION: CDNAS ENCODING PROTEINS CLOSELY RELATED
FILE REFERENCE: 0646/1A81BUST
CURRENT APPLICATION NUMBER: US/08/676,351C
CURRENT FILING DATE: 1996-09-12
EARLIER APPLICATION NUMBER: PCT/US95/00939
EARLIER FILING DATE: 1995-01-20
EARLIER APPLICATION NUMBER: US 08/185,360
EARLIER FILING DATE: 1994-01-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PaeSec for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2706
TYPE: DNA
ORGANISM: Rat
US-08-676-351-1

Alignment Scores:

Pred. No.: 6,89e-14 Length: 2706
 Score: 220.00 Matches: 75
 Percent Similarity: 42.90% Conservative: 64
 Best Local Similarity: 23.15% Mismatches: 117
 Query Match: 11.75% Indels: 68
 DB: 3 Gaps: 13

US-09-995-225B-16 (1-353) x US-08-676-351-1 (1-2706)

```

QY 7 HisLeuAlaAlaAsnSerLeuSerTrpSerProGlySerAlaCysGlyLeu 26
DB 200 CACCTGCTCTCCTCAATGCTAGTCACAGCCCTTCCTGCCC-----CTTGA 244
QY 27 Phe---ValProValTyrTyrSerLeuLeuCysLeu---GlyLeuProAla 44
DB 245 CTCAAGGTACACCATGGGGCTTACTTGGCTGTGTCATCGGGGGCTCTGGGAC 304
QY 45 IleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLeuSerSerTyrAsn 64
DB 305 TGCCCTGCTCATGTATGTCATC-----CTCAGGACACACAGATGAAGACGCTACCAAC 358
QY 65 ---TyrLeuLeuAlaLeuAlaAlaAlaAspIleLeuValIleuPheIleValPhe--- 82
DB 359 ATTTCACATTTTAACTGGGACCTGCTGATACCTGGCTTGTCTTACACAGCCCTTCAG 418
QY 83 ---ValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValProAsp 101
DB 419 GGCACAGACATCCCTACCTGGCTTGGCCATTGGGAAATGCACCTC-----TGC 466
QY 102 LysIleIleGluValLeuGlnPheSerSerIleHisThrSerIleTrpIleThrValPro 121
DB 467 AAGCTGTCTCATGCTGACCTACTACACATGTTACAGCACTTTTACTGACCGCC 526
QY 122 LeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSerTyr 141
DB 527 ATGACGCTAGACCGCTATGCTGCTATGCTGACACCTTACCTGATGTTCCGACA 586
QY 142 ProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSerIle 161
DB 587 TCCAGCAAGACCCGAGGTGTTATGTCGATATGGCCCTGCTTCAATGCTGTGTT 646
QY 162 ProTyr----- 163
DB 647 CCTGTTCCCATCATGGGTTCCAGACAACTGGAAGATGAAGATGAGTGGTGAG 706
QY 164 -----TyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThr 177
DB 707 ATCCCTGCCCCCTGAGGACTATGGGCCCT----- 736
QY 178 SerValHisIleValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSer 197
DB 737 ---GTTTGGCCATCTGCATCTTCTTTTCTTC-----ATCATCCCTGTGCTG 784
QY 198 IlePhePheIleLeuAsnSerIleIleValTyrLysLeuArg----- 211
DB 785 ATCATCTGCTGCTGCTAGAGCCTCATGATGACACATCTTGCGTGTCCGCTTCTTCA 844
QY 212 -----ArgLysSerAsnPheArgLeuArgLysTyrSerThrGlyThrThrAlaIle 229
DB 845 GGCCTCCCGGAGAGACCGAAACCTGGCG-----CGATCATCTGCAGCTG 889
QY 230 LeuPheThrIleThrSerIlePheAlaThrLeuTrpAlaProArgIleIleMetIleLeu 249
DB 890 GTGCTGTAGTGTGGCTGTGTTCTGTGGCTGCTGACGCTGTCAGAGTGTGTCCTG 949
QY 250 TyrHisLeuTyrGlyAlaProIleGlnAsnArgTrpLeuValHisIleMetSerAspIle 269
DB 950 GTTCAAGACAGTGGGTTCAGCCAGAGTAGAGACTGAGTGCATCTTCGCCG---TTC 1006
QY 270 AlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPheLeuTyrCysPheIleSer 289
  
```

```

DB 1007 TGCACAGCCCTGGGCTATGTCAACAGTGTCTCATGCCATTCTTATGCTTGAT 1066
QY 290 LysArgPheArg 293
DB 1067 GAGAACTTCAAG 1078
  
```

RESULT 14

US-09-016-434-1446

/ Sequence 1446, Application US/09016434

/ Patent No. 6500938

/ GENERAL INFORMATION:

/ APPLICANT: Janice Au-Young

/ APPLICANT: Jeffrey J. Seilhamer

/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

/ NUMBER OF INVENTIONS: PATHWAY GENE EXPRESSION

/ NUMBER OF SEQUENCES: 1490

/ CORRESPONDENCE ADDRESS:

/ ADDRESS: INCYTE PHARMACEUTICALS, INC.

/ STREET: 3174 PORTER DRIVE

/ CITY: PALO ALTO

/ STATE: CALIFORNIA

/ COUNTRY: USA

/ ZIP: 94304

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/09/016,434

/ FILING DATE: HERewith

/ CLASSIFICATION:

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER:

/ FILING DATE:

/ CLASSIFICATION:

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Zeller, Karen J.

/ REGISTRATION NUMBER: 37,071

/ REFERENCE/DOCKET NUMBER: PA-0002 US

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (650) 855-0555

/ TELEFAX: (650) 845-4166

/ INFORMATION FOR SEQ ID NO: 1446:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 1317 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ IMMEDIATE SOURCE:

/ LIBRARY: GENBANK

/ CLONE: G693907

/ US-09-016-434-1446

Alignment Scores:

Pred. No.: 2.63e-14 Length: 1317
 Score: 219.50 Matches: 72
 Percent Similarity: 44.82% Conservative: 75
 Best Local Similarity: 21.95% Mismatches: 130
 Query Match: 11.73% Indels: 51
 DB: 4 Gaps: 11

US-09-995-225B-16 (1-353) x US-09-016-434-1446 (1-1317)

```

QY 19 ProGlySerAlaCysGlyLeuGlyPheValProVal---ValTyrTyrSerLeuLeu 37
DB 165 CCGGGAGACCGCGGGCGCGGCGCATGCTGCTATCCAGTCACTTACGCGCTGGTGC 224
QY 38 CysLeuGlyLeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArg 57
DB 225 CTGTGTGGGGCTGTGGGGAACGCCCTGTCAATCTTGTGATCTTCCG---TACGCCAAG 281
QY 58 ArgGlnLysSerSerTyrAsnTyrLeuLeuAlaIleAlaAspIleLeuValLeu 77
  
```



```
Db      602 ATGAGCGTAGACCGCTATGTGCTATCTGCCACCCCTATCCGGCCCTTGATGTTCCGACA 661
QY      142 ProAlaArgThrArgValIleValSerValTyrIleThrCysPheLeuThrSerIle 161
Db      662 TCCACCAAGCCCGAGCTGTATATGTGCATATGCGCCCTGGCTTCAGTGTGTGCTT 721
QY      162 ProTyr----- 163
Db      722 CCTGTGGCATCATGGGTTCACACACAGTGAAGATGAGATCGAGTGCCTGGTGAG 781
QY      164 -----TyrIleProAsnIleThrGluAspTyrIleSerThr 177
Db      782 ATCCCTGCCCTCAGACTATTTGGGGCCCT----- 811
QY      178 SerValHisHisValIleIleThrIleHisCysPheThrValTyrLeuValProCysSer 197
Db      812 ---GATTCGCCCATCTGCATCTCTCTTTTCTCTTC-----ATCATCCCTGTGCTG 859
QY      198 IlePhePheIleLeuAsnSerIleIleValTyrIleLeuArg----- 211
Db      860 ATCATCTCTGTCTGTCTACAGCCTCATGATTCGACGACTTCGTGGTGTCCGTGCTTCA 919
QY      212 -----ArgLysSerAsnPheArgLeuArgGlyTyrSerThrGlyIleThrAlaIle 229
Db      920 GGCTCCCGGAGAGACCGAACCCTGCGG-----CGTATCACTCGACTG 964
QY      230 LeuPheThrIleThrSerIlePheAlaThrLeuThrAlaProArgIleIleMetIleLeu 249
Db      965 GTGCTGTAGTGTGGCTGTGTGTGTGGGCTGCTGGACGCTGTGCAGGTGTGTGCTG 1024
QY      250 TyrHisLeuTyrGlyAlaProIleGlnAsnArgTyrLeuValHisIleMetSerAspIle 269
Db      1025 GTTCACAGACTGGGTGTTCAGCCAGGTAGAGACTGCAGTGCATCTCGCGC---TTC 1081
QY      270 AlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhePheLeuTyrCysPheIleSer 289
Db      1082 TGCACAGCCCTGGGCTATGTCAACAGTTGTCTCAATCCCATCTCTATGCTTTCCTGAT 1141
QY      290 LysArgPheArg 293
Db      1142 GAGAACTTCAG 1153
```

Search completed: December 14, 2004, 22:27:16
Job time : 92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:35:30 ; Search time 456 Seconds
(without alignments)
4268.142 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1 1 MHTTAHLAANSLSLMSWSPG.....CIKMLVYQYDKKGRKIKVSP 353

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.epool/US0995225.rnuc.14122004.110734.18062/app.query.fasta_1.519
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=cct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEARSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FAPOP=6 -Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1872	100.0	1059	16	US-10-450-590-1	Sequence 1, Appli
2	1872	100.0	1062	9	US-09-995-225-15	Sequence 15, Appli
3	1872	100.0	1062	10	US-09-995-225-15	Sequence 15, Appli
4	1872	100.0	1062	16	US-10-296-294A-3	Sequence 3, Appli
5	1872	100.0	1062	16	US-10-296-294A-4	Sequence 4, Appli
6	1872	100.0	1202	14	US-10-094-417-1	Sequence 1, Appli
7	1872	100.0	2117	17	US-10-779-104-1	Sequence 1, Appli
8	1868	99.8	1343	10	US-09-813-432-21	Sequence 21, Appli
9	1868	99.8	1343	15	US-10-174-364-21	Sequence 21, Appli
10	1868	99.8	1343	16	US-10-246-583-21	Sequence 21, Appli
11	1868	99.8	1343	17	US-10-689-832-21	Sequence 21, Appli
12	1857	99.2	1119	10	US-09-813-432-19	Sequence 19, Appli
13	1857	99.2	1119	15	US-10-174-364-19	Sequence 19, Appli
14	1857	99.2	1119	16	US-10-246-583-19	Sequence 19, Appli
15	1857	99.2	1119	17	US-10-689-832-19	Sequence 19, Appli
16	1857	99.2	1130	16	US-10-333-946-25	Sequence 25, Appli
17	1857	99.2	2198	15	US-10-314-076-1	Sequence 1, Appli
18	1844	98.5	1062	14	US-10-012-140-12	Sequence 12, Appli
19	1844	98.5	1526	14	US-10-012-140-10	Sequence 10, Appli
20	1841	98.3	1110	15	US-10-314-076-3	Sequence 3, Appli
21	1841	98.3	2189	14	US-10-219-834-6	Sequence 6, Appli
22	1756	93.8	1038	14	US-10-094-417-19	Sequence 19, Appli
23	1747	93.3	1062	16	US-10-450-590-9	Sequence 9, Appli
24	1643	87.8	957	15	US-10-174-364-84	Sequence 84, Appli
25	1643	87.8	957	16	US-10-246-583-84	Sequence 84, Appli
26	1638	87.5	1032	15	US-10-079-384-25	Sequence 25, Appli
27	1638	87.5	1032	16	US-10-450-590-5	Sequence 5, Appli
28	1638	87.5	1070	16	US-10-450-590-4	Sequence 4, Appli
29	1638	87.5	1158	16	US-10-343-650A-21	Sequence 21, Appli
30	1638	87.5	1826	16	US-10-450-590-6	Sequence 6, Appli
31	1471	78.6	795	15	US-09-791-932-6	Sequence 6, Appli
32	1317	70.4	930	10	US-09-791-932-3	Sequence 3, Appli
33	912	48.7	1466	16	US-10-328-916-4	Sequence 4, Appli
34	742	39.6	1125	9	US-09-995-225-5	Sequence 5, Appli
35	737	39.4	1125	10	US-09-995-225-5	Sequence 5, Appli
36	737	39.4	1127	14	US-10-094-417-13	Sequence 13, Appli
37	737	39.4	1330	13	US-10-011-147-1	Sequence 1, Appli
38	737	39.4	1793	16	US-10-328-916-3	Sequence 3, Appli
39	737	39.4	1962	16	US-10-333-946-29	Sequence 29, Appli
40	737	39.4	1125	14	US-10-012-140-15	Sequence 15, Appli
41	733	39.2	1719	14	US-10-012-140-13	Sequence 13, Appli
42	733	39.2	34118	15	US-10-017-161-1071	Sequence 1071, Ap
43	668	35.7	34118	15	US-10-292-788-909	Sequence 909, App
44	668	35.7	34118	15	US-10-292-788-909	Sequence 909, App
45	242	12.9	3364	15	US-10-253-983-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-10-450-590-1
Sequence 1, Application US/10450590
Publication No. US20040076985A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
FILE REFERENCE: LIO316 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/450,590
CURRENT FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/254,923
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: US 60/280,110
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: US 60/299,474
PRIOR FILING DATE: 2001-06-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1059
TYPE: DNA

! ORGANISM: Homo sapiens
US-10-450-590-1

Alignment Scores:

Pred. No.:	8,74e-178	Length:	1059
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	16	Gaps:	0

US-09-995-225b-16 (1-353) x US-10-450-590-1 (1-1059)

```

QY      1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20
DB      1 ATGAGAGACACGACGCGCCACTGCGACCAAGCTGCTGTGTGTGCTCCCGCGC 60
QY      21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB      61 TCGGCGCTGGCGGCTTGGTTCGTGCCGTGTCTACAGCCTCTTGGCTGCTCGT 120
QY      41 LeuProAlaSerLeuThrValValLeuLeuSerGlnLeuValAlaArgArgGlnGly 60
DB      121 TTACCGCAAAATATCTTGACAGATCATCTCTCCAGCTGGCGCAAGAGAGAGAA 180
QY      61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspLeuValLeuValLeuPhePhe 80
DB      181 TCCTCCGACAACTATCTCTGGCACTGCGCTGCGCGACACTTGTGGCTCTTTTCATA 240
QY      81 ValPheValAspPheLeuLeuGluAspPheLeuAsnMetGlnMetProGlnValPro 100
DB      241 GTGTTTGTGACCTTCTGTGGAAAGATTTCATCTTGAACATGACATGCTGAGTCCC 300
QY      101 AspLeuLeuLeuGluValLeuGluPheSerSerLeuHisThrSerLeuTrpLeuThrVal 120
DB      301 GACAAATCATTAAGTGTCTGGAATTCATTCATCAACACCTCCATATGATGATTAATCTGA 360
QY      121 ProLeuThrLeuAspArgTyrLeuAlaValCysHisProLeuLeuTyrHisThrValSer 140
DB      361 CCGTTAACCATTCAGAGTATATCGCTGCGCACCGCTCAAGTACACACGAGTCTCA 420
QY      141 TyrProAlaArgThrArgValValLeuValSerValTyrLeuThrCysPheLeuThrSer 160
DB      421 TACCCGCGCGGACCGGAAAGCATTTGATGTTTACATCACCCTGCTCTGACGACG 480
QY      161 IleProTyrTyrTrpProAsnLeuTrpThrGluAspTyrLeuSerThrSerValHis 180
DB      481 ATCCCTATTTACGTGGTGGCCCAACATCTGACTGAAGACTACATACGACCTCTGTGAT 540
QY      181 HisValLeuLeuLeuTrpLeuHisCysPheThrValTyrLeuValProCysSerLeuPhePhe 200
DB      541 CACGCTCATCTGATCCATCCATGCTTCAACGCTCTACCTGCTCCATCTTCTTC 600
QY      201 IleLeuAsnSerLeuLeuValTyrLeuValArgGlySerAsnPheArgLeuArgGly 220
DB      601 ATCTTGAATCATATCTTGTGTACAGCTCAGAGAGAGAGAGCAATTTTCGTCGCTGAC 660
QY      221 TyrSerThrGlyLeuThrThrAlaIleLeuPheThrLeuThrSerLeuPheAlaThrLeu 240
DB      661 TACTCCAGCGGAGAACACCGCCCATCTTGTTCACCATTTACCTCATCTTGGCCACACTT 720
QY      241 TrpAlaProArgGlyIleMetLeuLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
DB      721 TGGGCGCGCGCATATCATATGTTCTTTACCACTCTTATGGGCGCGCCATCCAGAACGCG 780
QY      261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
DB      781 TGGCGTGAACACATCATCTCCGACCTTCCAAACATGCTAGCCCTTGAAGACAGCCATC 840
QY      281 AsnPhePheLeuTyrCysPheLeuSerLeuValArgPheArgThrMetAlaAlaAlaThrLeu 300
DB      841 AACTTCTTCTCTACTGCTCATCAGCAAGCGGTTCCGACATAGGACGCGCCACGCTC 900

```

```

QY      301 TyrAlaPhePheLeuCysGlnGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB      901 AAGGCTTTCTCAAGTGCACAGAGCAACCTGTACAGTTCTTACACCAATCATTAATCTTCC 960
QY      321 IleThrSerSerProTrpTrpLeuSerProAlaAsnSerHisCysIleLeuMetLeuValTyr 340
DB      961 ATTAACAAGTAGAGCCCTTGATTCGTGCGCGCAACTACACTGCATCAAGATGCTGTGTAC 1020
QY      341 GlnTyrAspLeuAsnGlyLeuSerProIleLeuValSerPro 353
DB      1021 CAGTATGACAAATAATGAAAACTTATTAAGATATCCCG 1059

```

RESULT 2

US-09-995-225-15

Sequence 15, Application US/09995225

Publication No. US20020193584A1

GENERAL INFORMATION:

APPLICANT: Chen, Ruoping

APPLICANT: Chu, Zhi Liang

APPLICANT: Dang, Huang T.

APPLICANT: Lowitz, Kevin P.

APPLICANT: Pride, Cameron

TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G

TITLE OF INVENTION: Receptors

FILE REFERENCE: AREN-0308

CURRENT APPLICATION NUMBER: US/09/995, 225

CURRENT FILING DATE: 2001-11-26

PRIOR APPLICATION NUMBER: 09/170,496

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: PCT/US99/23938

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253,404

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/255,366

PRIOR FILING DATE: 2000-12-12

PRIOR APPLICATION NUMBER: 60/270,286

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,365

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/270,266

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,032

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,358

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,356

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/290,917

PRIOR FILING DATE: 2001-05-14

PRIOR APPLICATION NUMBER: 60/309,208

PRIOR FILING DATE: 2001-07-31

NUMBER OF SEQ ID NOS: 67

SOFTWARE: PatentIn version 3.1

SEQ ID NO 15

LENGTH: 1062

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: No. US20020193584A1 Sequence

US-09-995-225-15

Alignment Scores:

Pred. No.:	8,74e-178	Length:	1062
Score:	1872.00	Matches:	353
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-995-225b-16 (1-353) x US-09-995-225-15 (1-1062)

```

QY      1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20

```


Db 1 ATGGAGACACGACGCCACCTCGGAGCCACAGCTCGCTGTGTGTGTCCTCCCGGC 60
Qy 21 SerAlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCyLeuGly 40
Db 61 TCGGCGTGGGCTTGGGTTTGTGCGCGGTGTACTACAGCTCTTGCTGCTCGGT 120
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db 121 TTAACGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGGTGGCAAGACAGAG 180
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 181 TCCTCTTACAACATATCTTGGCACTCGCTGCGGACATCTTGGCTCTTTTCATA 240
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 241 GTGTTGTGTGACTCTGTTGGAAAGATTCATCTTGAAACATGACAGATGCTCAGGTC 300
Qy 101 AspPheIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrrIleThrVal 120
Db 301 GACAAATCATATAGAGTGTGGAATTCATCCATCCACCTCCATATGATTTACTGTA 360
Qy 121 ProLeuThrIleAspArgTyrIleAlaValAlaCyHisProLeuLysTyrHisThrValSer 140
Db 361 CGTTTAAACATTGACAGGTATATCGGTGTGCGCACCGGCTCAAGTACCAAGGCTCTCA 420
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 421 TACCCGACCCGCAACCCGAAAGCATGTGTAAAGTTTAACTACCTGCTTCGACACAG 480
Qy 161 IleProTyrTyrTrrTrrProAsnIleTrrThrGluAspTrrIleSerThrSerValHis 180
Db 481 ATCCCTATATACGTGGGCGCCCAACATCTGGACGTGAAGACTACATCAGCACTCTGTCAT 540
Qy 181 HisValLeuIleTrrPheHisCysPheThrValTyrIleValProCysSerIlePhePhe 200
Db 541 CAGCTCTCATCTGGATCCACTGCTTCAACGCTGTACTGGTGGCCGTCTCATCTTCTTC 600
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgLys 220
Db 601 ATCTTGAACTCATCATTTGTGTCAAGCTCAGAGAGAGAGCAATTTTCCTCCCTGCGC 660
Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleHisSerIlePheAlaThrIleu 240
Db 661 TACTCCAGCGGAGAGACACCGCCATCTGTGTACCACTTACCTCCATCTTGGCACACTT 720
Qy 241 TrrAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 721 TGGGCCCCCGCATCATCATGATTCCTTACCACTCTATGGGCGCCCATCCAGAACCGC 780
Qy 261 TrrLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 781 TGGCTGTGTGACATCATGTCTCGCATTTGCCAATCATCTGACCTTGAACACAGCCATC 840
Qy 281 AsnPhePheLeuLysCysPheIleSerLysArgPheArgThrMetIleAlaAlaThrIleu 300
Db 841 AACTTTCTTCTTACTAGCTTCATCAGACAGCGGTTCGCAACATGGAGCGCCAGCGCTC 900
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 901 AAGGCTTTCTTCAAGTCCAGAGAACCGTACAGTTCACACCAATCATTACTTTTTC 960
Qy 321 IleHisSerSerProTrrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 961 ATTAACAAGTAGCCCTCGGATCTCGCGGCAAACTCACACTGATCAAGAAGTGTGTGTAC 1020
Qy 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
Db 1021 CAGTATGACAAATAATGAAAACTATTAAGTATCCCGC 1059

RESULT 3
US-09-995-225-15
; Sequence 15, Application US/09995225

Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Kuoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9el Sequence
US-09-995-225-15
Alignment Scores:
Pred. No.: 8,74e-178 Length: 1062
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-995-225B-16 (1-353) x US-09-995-225-15 (1-1062)
Qy 1 MetGlnHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrrTrrSerProGly 20
Db 1 ATGGAGACACGACGCCACCTCGGAGCCACAGCTCGCTGTGTGTGTCCTCCCGGC 60
Qy 21 SerAlaCyGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCyLeuGly 40
Db 61 TCGGCGTGGGCTTGGGTTTGTGCGCGGTGTACTACAGCTCTTGCTGCTCGGT 120
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db 121 TTAACGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGGTGGCAAGACAGAG 180
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 181 TCCTCTTACAACATATCTTGGCACTCGCTGCGGACATCTTGGCTCTTTTCATA 240
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100


```

QY 28 AaNPhePheLeuYrCySPheIleSerLyAaPhePheqThrMetAlaIaIaIaThrLeu 300
Db 841 AACTTCTTCTCTACTGCTTCAATCAGCAAGCGGTTCGCGACCAAGCGCCGACGCTC 900
QY 301 LySaIaPhePheLySCySgInLySgInProValGInPheTYrThraSnnHiaSPheSer 320
Db 901 AAGGCTTTCTTCAAGTCCAGAAAGCAACCTGTACAGTTCTTACCAATCAATCACTTTTCC 960
QY 321 ILeThrSerSerProTrpIleSerProAlaSnnSerHicSvIleYMetLeuValTYr 340
Db 961 ATAAACAAGTAGCCCTCGATCTCGCCGCGCAAACTCACATGCAATCAAGATCGTGTATAC 1020
QY 341 GInTYrAspLyAaSnGlyLySProlleLySProlleValSerPro 353
Db 1021 CAGTATGACAAATAATGAAAACCTATTAATAAGTATCCCG 1059

RESULT 5
US-10-296-294A-4
; Sequence 4, Application US/10296294A
; Publication No. US20040029224A1
; GENERAL INFORMATION:
; APPLICANT: TERAQ, Yasuko
; APPLICANT: MATSUI, Hideki
; APPLICANT: SHIMIZU, Yasushi
; TITLE OF INVENTION: No. US20040029224A1el G Protein-Coupled Receptor and its DNA
; FILE REFERENCE: 2734 USOP
; CURRENT APPLICATION NUMBER: US/10/296, 294A
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: PCT/JP01/04643
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: JP 2000-170446
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: JP 2000-194926
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Human
US-10-296-294A-4

Alignment Scores:
Pred. No.: 8,74e-178 Length: 1062
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-995-225B-16 (1-353) x US-10-296-294A-4 (1-1062)
QY 1 MetGluHleThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTYrTrpSerProGly 20
Db 1 ATGGAGCACAAGCAGCCCACTCGCGAGCCAAACAGCTCGCTGTGGTGTGCTCCCGCGC 60
QY 21 SerAlaCySgIyLeuGlyPheValProValValTYrTYrSerLeuLeuLeuCyaleuGly 40
Db 61 TCGGCTGGGCGCTTGAGTTCGTGCGCCGCGTGTCTACTACAGCCTTGTGTGCTCGGT 120
QY 41 LeuProAlaSnnIleLeuThrValIleIleLeuSerGlnLeuValAlaAgaArgGlnLyS 60
Db 121 TTACAGCAAAATATCTTGACAGTATCTCTCCCAAGCTGGGCGCAAGACAGAG 180
QY 61 SerSerTYrAsnTYrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 181 TCCTCTTCAACTATCTCTTGGCATCTCGCTGCGCAGCAATCTTGATCTCTTTTATA 240
QY 81 ValPheValAspPheLeuLeuGlnAspPheIleLeuSnnMetGlnMetProGlnValPro 100
Db 241 GGTGTGTGAGACTTCTGTGGAAGATTCATCTTGAAACATGACGATGCTCAGTCCC 300
QY 101 AspLyIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120

```

Db	301	GACAAAGATCATAGAAGAGTGGGAATTTGCATTCATCCATCCACACCTCCATATGAGATTACTGTA	360
Qy	121	ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuValTyrHisThValSer	140
Db	361	CCGTTAACATATGACAGGTATATGCTGTCGCCACCGCTCAAGTACACACAGCTGTCA	420
Qy	141	TyrProAlaAspThrArgValIleValSerAlaTyrIleThrCysPheLeuThrSer	160
Db	421	TACCAAGCCGACCCGGAAGTCAATGTAAGTATTATCATCACCCTGCTTCTTAACACAGC	480
Qy	161	IleProTyrTyrTrpTrpProAsnIleTrpHisIuAspTyrIleSerThrSerValHis	180
Db	481	ATCCCTTATTTCTGTGGCCCAACATCTGGACCTAAGACTATCATCAGACCTGTGTGAT	540
Qy	181	HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe	200
Db	541	CAGCTCCATCTGATGCATCCATGCTTCAACCTTAACCTGGAGCCCTGCTTCATCTTCTTC	600
Qy	201	IleLeuAsnSerIleIleValTyrLeuLeuArgArgIysSerAsnPheArgLeuArgGly	220
Db	601	ATCTTGAACTCAATCATTTGTGTAAAGTCAAGGAGGAGAAAGCAATTTCTGCTCCGTGC	660
Qy	221	TyrSerThrGlyTyrThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu	240
Db	661	TACTTCACGGGGAAGACACCGGCATCTGTTCACCATTACTCATCTTGTGCCACATT	720
Qy	241	TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg	260
Db	721	TGGGCCCCCGCATCATCATCATATTCTTTACCACTTATGGGGGCCCATCCAGAACCGC	780
Qy	261	TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle	280
Db	781	TGGGTGGGCATCATCATGTCCGACATTCGCCAACATGCTTACGCCCTTGGAACACAGCCATC	840
Qy	281	AsnPhePheLeuTyrCysPheIleSerIysArgPheArgThrMetAlaAlaAlaThrLeu	300
Db	841	AACCTTCTTCTCTACTGCTTCATACGACAGCGGTCTCCGACCATGGACCGGCCACGCTC	900
Qy	301	LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer	320
Db	901	AAGCTTCTTCTTCAAGTGCAGAAACCAACCTGTACAGTTCTACCAATCATTACTTTTCC	960
Qy	321	IleHisSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr	340
Db	961	ATAACAAGTACCCCTGTGATCTCCGCGCAAACTCACACTGCATCAAGATGCTGTGTAC	1020
Qy	341	GlnTyrAspLysAsnGlyLysProIleLysValSerPro	353
Db	1021	CAGTATGACAAATAATGAAAAACCTATAAAGTATCCCCG	1059

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (68)..(1129)
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
; US-10-094-417-1

Alignment Scores:
Pred. No.: 1.05e-177 Length: 1202
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-995-225b-16 (1-353) x US-10-094-417-1 (1-1202)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20
DB 68 ATGGAGACACGACGCCACCTCGACGCCAACAGCTCGCTGCTGGAGTCCCCCGGC 127
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB 128 TCGGCTGCGGCTGGGTTTCGTCGCGGTCTACTACAGCCTTGCTGTCCTCGGT 187
QY 41 LeuProAlaAsnLeuLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
DB 188 TTACCAAGCAATATCTTGACAGATGATCTCTCCAGTGGTGGCAAGAACAGAAAG 247
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
DB 248 TCCTCTACACATATCTTGACCTCGCTGCTGCGGACATCTGCTCTTTTTCATA 307
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnVal 100
DB 308 GTGTGTGTGACTTCTGTTGGAAAGATTCACTTGAACATGAGATGCTCAGGTCCC 367
QY 101 AspLeuIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
DB 368 GACAGATCATAGAAAGTCTGGAATCTCATCTCATCCACACCTCCATATGATGTA 427
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLeuTyrHisThrValSer 140
DB 428 CCGTTAACCATGACAGATATGCTGCTGCGACCCGCTCAAGTACACACGGCTCTCA 487
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 488 TACCCAGCCCGACCCGGAAGTCATGTAGTGTTCATCATCACCTGCTTCTGACACAGC 547
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
DB 548 ATCCCTTATTAATGAGTGGCCCAACATCTGACTGAAAGTACATACACCTCTTGAT 607
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
DB 608 CACGCTTCATCTGATCCACATGCTTCAACCGTCTACCTGCTGCTTCCATCTTCTTC 667
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgGlySerAsnPheArgLeuArgLys 220
DB 668 ATCTTGAATCATATCTGTGTGACAGCTCAGAGAGAAAGCAATTTTCTGCTGAGC 727
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 728 TACTCAGCGGGAACACACCGCCATCTTGTTACCATTAATCTTCATCTTGCCACACTT 787
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
DB 788 TGGGCCCCCGCATCATGATTTCTTTAACCACTCTATGGGGCCCATCAGAACCGC 847
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
```

```

DB 848 TGGCTGTACACATCATGTCCACATTCGCCAACATGTCAGCCCTTCTGAAACACAGCCATC 907
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
DB 908 AACTTCTCTCTACAGCTTCAATCAGCAAGCGGTTCCGACCATGCGACCCCGACGCTC 967
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 968 AAGGCTTCTTCAAGTCCAGAACCAACCTGTACAGTTCTACACCAATCATTAATCTTTC 1027
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysAlaLysMetLeuValTyr 340
DB 1028 ATPACAGTAGCCCTTGATCTCGCCGCAACTACACTGATCATCAAGTCTGGGTGAC 1087
QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
DB 1088 CAGTATGACAAATAATGAAAACCTATTAAGATATCCCCG 1126
```

RESULT 7

```

US-10-779-104-1
; Sequence 1, Application US/107799104
; Publication No. US20040161799A1
; GENERAL INFORMATION:
; APPLICANT: Andrew J. Murphy
; APPLICANT: Susan Croll-Kalish
; TITLE OF INVENTION: KOR3-like-Proteins and Methods of Modulating KOR3J-Mediated Activi
; FILE REFERENCE: REG 1000A
; CURRENT APPLICATION NUMBER: US/10/779,104
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: 60/447,447
; PRIOR FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 60/495,577
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-779-104-1
```

Alignment Scores:

```

Pred. No.: 2.48e-177 Length: 2117
Score: 1872.00 Matches: 353
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
```

US-09-995-225b-16 (1-353) x US-10-779-104-1 (1-2117)

```

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpTrpSerProGly 20
DB 280 ATGGAGACACGACGCCACCTCGACGCCAACAGCTCGCTGCTGGAGTCCCCCGGC 339
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
DB 340 TCGGCTGCGGCTGGGTTTCGTCGCGGTCTACTACAGCCTTGCTGTCCTCGGT 399
QY 41 LeuProAlaAsnLeuLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
DB 400 TTACCAAGCAATATCTTGACAGTATCATCTCTCCAGCTGAGTGGCAAGACAGAG 459
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePhe 80
DB 460 TCTCTTCAACATATCTTGAGCACTGCTGCGCCACATCTTGGTCCCTTTTCATA 519
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 520 GTGTGTGTGACTTCTGTTGGAAGATTTCACTTGAACATGAGATGCTCAGGTCCC 579
QY 101 AspLeuIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
```

```

Db      580 GACAGATCATGAGAGTGTGAAATTCATCCATCCACACCTCCATATGATTAAGTGA 639
Qy      121 Proleuthr11eapargtr11ealava1CysH1sProleu1ytr1H1sthrValSer 140
Db      640 CCGTAAACCATTAACAGATATATCGCTGTGCAACCCGCTCAAGTACCAACAGAGCTCA 639
Qy      141 TyrPro1a1argthr1arg1yval11e1a1SerVal1Tr11e1thrCysPhe1euthrSer 160
Db      700 TACCCAGCCCGCAACCCGGAAGTCACTTGTAAAGTGTTAATCACTGCTCTGACACAGC 759
Qy      161 11ePro1yTr1Tr1P1Proa1n1e1Tr1P1Th1G1uA1p1yTr11e1Ser1H1s 180
Db      760 ATCCCTATATAGTGTGGCCCAACATCTGACATGAAGACTACATACGACACTCTGTGCAT 819
Qy      181 H1eVal1leu11e1Tr1P11e1H1sCysPhe1Th1Val1Ty1r1e1u1Val1Pro1Cys1er11e1Phe 200
Db      820 CACGCTCTATCTGAGATCCACCTGCTTCAACGCTACCTGTGCTGCTCCATCTTCTTC 879
Qy      201 11e1e1uA1n1Ser11e11e1Val1Ty1r1y1b1e1u1Arg1y1Ser1a1n1Phe1arg1e1u1Arg1y 220
Db      880 ATCTTAATCACTCAATCTTGTGTCAAGCTCAAGAGAGAGCAATTTTCGTCCTCGTGC 939
Qy      221 TyrSer1Th1G1y1y1Th1Th1a11e1e1u1Phe1Th11e1Th1Ser11e1Phe1a1Th1r1e1u 240
Db      940 TACTCCACGGGGAAGACACCGCCATCTTGTACCATTAACCTCCATCTTGTGCACACTT 999
Qy      241 Tr1P1a1P1ro1Arg11e11e1e1e11e1e1u1Ty1H1s1e1u1Ty1G1y1A1P1ro11e1G1n1a1n1Arg 260
Db      1000 TGGGCCCCCGCATCATCATGATTCCTTTCACACCTCATGCGGCGCCATCCAGAAACCGC 1059
Qy      261 Tr1P1e1u1Val1H1s11e1e1Ser1a1P11e1a1a1n1Met1e1u1a11e1u1e1u1n1Th1r1a11e 280
Db      1060 TGCGTGTGACATCATGTCCTCCATTCGCCAATGCAACATGCAACCTTCTGAAACAGACCATC 1119
Qy      281 Asn1Phe1Phe1e1u1Ty1CysPhe11e1Ser1y1a1r1Phe1a1Th1r1e1a1a1a1a1Th1r1e1u 300
Db      1120 AACTTCTTCTCTCAATGCTTCAACAGAGCGCTTCCGACCATGAGAGCGGCGACGCTC 1179
Qy      301 Lys1Ala1Phe1Phe1y1CysG1n1y1G1n1P1ro1Val1G1n1Phe1Ty1Th1r1a1n1H1s1a1n1Phe1Ser 320
Db      1180 AAGGCTTCTTCAAGTGCAGAGCAACCTGTAACAGTTCACACATCACTTAATCTTTC 1239
Qy      321 11e1Th1Ser1Ser1Pro1Tr1P11e1Ser1Pro1a1a1n1Ser1H1sCys11e11e1y1Met1e1u1Val1Ty1r 340
Db      1240 ATTAACAAGTAGCCCTGAGATCTGCGGCAACCTCACTGCAATCAAGATGTGTGTAC 1299
Qy      341 Gln1Ty1Asp1y1Asn1G1y1y1P1ro11e1y1Val1Ser1Pro 353
Db      1300 CAGTATGACAAATAATGAAAACTTAATAAGTATCCCGC 1338

RESULT 8
US-09-813-432-21
; Sequence 21, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Majumder, Kamud
; APPLICANT: Spaderma, Steven K
; APPLICANT: Smithson, Glenda
; APPLICANT: Mezes, Peter S
; APPLICANT: Vermet, Corine A. M.
; TITLE OR INVENTION: No. US20030148485A1 Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729
; CURRENT APPLICATION NUMBER: US/09/813,432
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22

```

```

; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-813-432-21

Alignment Scores:
Pred. No.: 3,14e-177 Length: 1343
Score: 1968.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 10 Gaps: 0

US-09-995-225b-16 (1-353) x US-09-813-432-21 (1-1343)
Qy      1 MetG1uH1s1Th1H1e1a1H1s11e1u1a1a1a1n1Ser1e1u1Ser1Tr1P1Ser1Pro1y 20
Db      2 ATGAGACACACGACCCACCTCGACGCAACAGCTCGCTGTGTGTCCTCCCGGC 61
Qy      21 Ser1Ala1CysG1y1Leng1y1Phe1Val1P1ro1Val1Ty1Tr1y1Ser1e1u1e1u1Cys1e1u1G1y 40
Db      62 TCGGCTGCGGCTGGGTTCGTCGCGGTGTCTACTAAGCTTGTGTCCTGCTCGGT 121
Qy      41 Leu1Pro1a1a1n11e1e1u1Th1r1Val11e11e1e1u1Ser1G1n1e1u1Val1a1a1r1Arg1n1y1 60
Db      122 TTAACGACAAATATCTTGACATGATCATCTCTCCACCTGTGTCAGAGACAGAG 181
Qy      61 Ser1Ser1Ty1Asn1Ty1r1e1u1a1a1e1u1a1a1a1a1a1P1r1e1u1Val1e1u1Phe1Phe1 80
Db      182 TCTCTCAATATCTTGTGACATCGCTGCGGACATCATCTGTGCTCTTTTCATA 241
Qy      81 Val1Phe1Val1Asp1Phe1e1u1e1u1G1uA1p1H1e11e1u1a1n1Met1G1n1e1r1Pro1G1n1Val1Pro 100
Db      242 GTGTTGTGAGCTTCTCTGTGAGAGATTCATCTTGAAATGCAATGCAATGCTCAAGTCC 301
Qy      101 Asp1y11e11e1G1u1Val1e1u1G1uPhe1Ser11e1H1s1Th1Ser11e1Tr1P11e1Th1r1Val 120
Db      302 GACAGATCATGAGAGTGTGATATCTGATCATCAACACCTCCATATGATTAAGTGA 361
Qy      121 Proleuthr11eapargtr11ealava1CysH1sProleu1ytr1H1sthrValSer 140
Db      362 CCGTAAACCATTAACAGATATATCGCTGTGCAACCCGCTCAAGTACCAACAGAGCTCA 421
Qy      141 TyrPro1a1argthr1arg1yval11e1a1SerVal1Tr11e1thrCysPhe1euthrSer 160
Db      422 TACCCAGCCCGCAACCCGGAAGTCACTTGTAAAGTGTTAATCACTGCTTCTGACACAGC 481
Qy      161 11ePro1yTr1Tr1P1Proa1n1e1Tr1P1Th1G1uA1p1yTr11e1Ser1H1s 180
Db      482 ATCCCTATATAGTGTGGCCCAACATCTGACATGAAGACTACATACGACACTCTGTGCAT 541
Qy      181 H1eVal1leu11e1Tr1P11e1H1sCysPhe1Th1Val1Ty1r1e1u1Val1Pro1Cys1er11e1Phe 200
Db      542 CACGCTCTATCTGAGATCCACCTGCTTCAACGCTACCTGTGCTGCTCCATCTTCTTC 601
Qy      201 11e1e1uA1n1Ser11e11e1Val1Ty1r1y1b1e1u1Arg1y1Ser1a1n1Phe1arg1e1u1Arg1y 220

```

```
Db      602 ATCTGAACATCATCTTGTGTACAGCTCAGAGAGAAAGCAATTTCTGTCCTCGGTGAC 661
Qy      221 TysSerThGlyLysThrThraAlaIleuPheThrIleThSerIlePheAlaThrIleu 240
Db      662 TACTCCACGGGAGAGACACCGCATCTGTTCAACATTACTCCATCTTGCCACACTT 721
Qy      241 TTPAlaProArgIleIleMetIleLeuTyrlsLeuTyrgIyAlaProIleGlnAsnArg 260
Db      722 TGGGCCCCCGCATCATCATGATCTTTCATCAACCTCTATGGGGCCGCCATCCAGAACCGC 781
Qy      261 TTPLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      782 TGGCTGTACACATCATGTCGACATTCGCAACATGCTAAGCCTTCTTAACACAGCACATC 841
Qy      281 AsnPhePheLeuTyrcysPheIleSerIleSarIleArgPheAlaIleAlaIleThrIleu 300
Db      842 AACTCTTCTCTTACTGCTTCACTACGACGAGGCGTTCGACCATGGCAGCGCCGACGCTC 901
Qy      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrlsThrAsnHisAsnPheSer 320
Db      902 AAGGCTTCTTCAAGTGCACAGAGCAACTGTACAGTTCTACACCAATCATTAATCTTTCC 961
Qy      321 IleThrSerSerProThrIleSerProAlaAsnSerHisCysIleIleYsMetLeuValTyrls 340
Db      962 ATTAACAAGTACGCCCTGATCTCCGCGCAAACTCACACTGCATCAAGATGCTGTGTAC 1021
Qy      341 GlnTyraAspLysAsnGlyLysProIleLysValSerPro 353
Db      1022 CAGTATGACAAATAATGGAAAACCTATAAAGTATCCCCG 1060
```

RESULT 9

```
US-10-174-364-21
: Sequence 21, Application US/10174364
: Publication No. US20030216308A1
: GENERAL INFORMATION:
: APPLICANT: Anderson et al.
: TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: 15966-729CIP2
: CURRENT APPLICATION NUMBER: US/10/174,364
: PRIOR FILING DATE: 2002-06-17
: PRIOR APPLICATION NUMBER: 60/190,835
: PRIOR FILING DATE: 2000-03-20
: PRIOR APPLICATION NUMBER: 60/190,768
: PRIOR FILING DATE: 2000-03-20
: PRIOR APPLICATION NUMBER: 60/190,972
: PRIOR FILING DATE: 2000-03-22
: PRIOR APPLICATION NUMBER: 60/191,199
: PRIOR FILING DATE: 2000-03-22
: PRIOR APPLICATION NUMBER: 60/191,947
: PRIOR FILING DATE: 2000-03-24
: PRIOR APPLICATION NUMBER: 60/192,665
: PRIOR FILING DATE: 2000-03-28
: PRIOR APPLICATION NUMBER: 60/192,657
: PRIOR FILING DATE: 2000-03-28
: PRIOR APPLICATION NUMBER: 60/192,984
: PRIOR FILING DATE: 2000-03-28
: PRIOR APPLICATION NUMBER: 60/192,664
: PRIOR FILING DATE: 2000-03-28
: PRIOR APPLICATION NUMBER: 60/192,836
: PRIOR FILING DATE: 2000-03-29
: REMAINING Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 128
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 21
: LENGTH: 1343
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (2) ..(1060)
US-10-174-364-21
```

Alignment Scores:

```
Pred. No.: 3,14e-177 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 15 Gaps: 0
```

US-09-995-225b-16 (1-353) x US-10-174-364-21 (1-1343)

```
Qy      1 MetGlnHisThrHisAlaIleuAlaAsnSerLeuSerTrpTrpSerProGly 20
Db      2 ATGAGACACACGACCGCCACTCCGACGCAACACTGCTGTCTTGAGTGATCCCCCGGC 61
Qy      21 SerAlaCysGlyLeuGlyPheValProValValTyrlsSerLeuLeuCysLeuGly 40
Db      62 TGGGCTGGCGGCTTGGGTTTCGTGCCCGGTGTACTACAGCTTCTTGCTGTGCTGGGT 121
Qy      41 LeuProAlaAsnIleLeuThrValIleIleuSerGlnLeuValAlaArgArgGlnLys 60
Db      122 TTACCAAGAAATATCTTACAGTATCATCTCTCCAGCTGTGGCAAGAGACAGAAAG 181
Qy      61 SerSerTyraTyrlsLeuLeuAlaLeuAlaAlaAspIleLeuValIleuPhePheIle 80
Db      182 TCCTCTCAACTATCTCTTGGCACTGCTGCTGCCGACATCTTGCTCTTTTCATA 241
Qy      81 ValPheValAspPheLeuLeuGlnAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      242 GTGTTTGTGGACTTCCTGCTTGAGAAATTTCACTTGAACATGACAGATGCTTCAGTCCC 301
Qy      101 AspIleIleIleGluValIleuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db      302 GACAGATCATATGAAAGGCTGGATTCATCTCATCTCACACCTCCATATGATTACTGTA 361
Qy      121 ProLeuThrIleAspArgTyrlsIleAlaValCysHisProLeuLysTyrlsIleThrValSer 140
Db      362 CCGTTAACCTTACACAGTATATCACTGTGCACCCGCTCAAGTACACACGCGTCTCA 421
Qy      141 TyrlsProAlaArgThrArgLysValIleValSerValTyrlsIleThrCysPheLeuThrSer 160
Db      422 TACCCAGCCCGCACCCCGAAAGTCAATGTAAAGTTTAACTACACCTGCTTCCGACCA 481
Qy      161 IleProTyrlsTrpTrpProAsnIleTrpThGlnAspTyrlsSerThrSerValHis 180
Db      482 ATCCCATTAATCTGTGGGCCCAACATCTGACTGAAGATCAATCAGCACCTCTGGCAT 541
Qy      181 HisValIleuIleTrpIleHisCysPheThrValTyrlsLeuValProCysSerIlePhePhe 200
Db      542 CACGTCCTCATCTGATCCACAGCTTCACCGTCTACCTGTGCTCCCTGCTCCATCTTCTTC 601
Qy      201 IleLeuAsnSerIleIleValTyrlsLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db      602 ATCTTAACATCATCTTGTGTACAGCTCAGAGAGAAAGCAATTTCTGTCCTCGTGGC 661
Qy      221 TysSerThGlyLysThrThraAlaIleuPheThrIleThSerIlePheAlaThrIleu 240
Db      662 TACTCCACGGGAGAGACACCGCATCTGTTCAACATTACTCCATCTTGCCACACTT 721
Qy      241 TTPAlaProArgIleIleMetIleLeuTyrlsLeuTyrgIyAlaProIleGlnAsnArg 260
Db      722 TGGGCCCCCGCATCATCATGATCTTTCATCAACCTCTATGGGGCCGCCATCCAGAACCGC 781
Qy      261 TTPLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      782 TGGCTGTACACATCATGTCGACATTCGCAACATGCTAAGCCTTCTTAACACAGCACATC 841
Qy      281 AsnPhePheLeuTyrcysPheIleSerIleSarIleArgPheAlaIleAlaIleThrIleu 300
Db      842 AACTCTTCTCTTACTGCTTCACTACGACGAGGCGTTCGACCATGGCAGCGCCGACGCTC 901
Qy      301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrlsThrAsnHisAsnPheSer 320
Db      902 AAGGCTTCTTCAAGTGCACAGAGCAACTGTACAGTTCTTAACCAATCATTAATCTTTCC 961
```

QY 321 IleThrSerSerProTrrPleSerProAlaSerSerHisCyBilelyMetLeuValTyr 340
Db 962 ATACAGAGAGAGCCCTGGATCTCGCCGCAACTCAGCATCAGAGATGGTGATC 1021
QY 341 GlnTyrAspLysAsnGlyLysProIleLeuValSerPro 353
Db 1022 CAGTATGACAAATAATGAGAAACCTATTAAGTATCCCCG 1060

RESULT 10
US-10-246-583-21
; Sequence 21, Application US/10246583
; Publication No. US20040058862A1
; GENERAL INFORMATION:
; APPLICANT: Majumder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1060)
US-10-246-583-21

Alignment Scores:
Pred. No.: 3,14e-177 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 16 Gaps: 0

US-09-995-225b-16 (1-353) x US-10-246-583-21 (1-1343)

QY 1 MetGluHisThrIleValIleAlaSerSerLeuSerTrrPTrSerProGly 20
Db 2 ATGAGACACAGCAGCCCACTCGCAGCCAGCCAGCTCGTGTGGTGATCCCCGGC 61
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
Db 62 TCGGCTGGGGCTTGGGTTTCTGCGCCGCTGCTTACTACAGCTCTTGGCTGCTCGGT 121
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db 122 TTACGACGAAATATCTTGAAGATGATCATCTCTCCAGCTGGTGCAAGAGACAGAG 181

QY 61 SerSerTyrAsnThrLeuLeuAlaAlaAlaSerIleLeuValLeuPhePheIle 80
Db 182 TCCTCCACACATATCTTGGACATCGCTGGCCGACATCTTGCTCTTTTCATA 241
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnVal 100
Db 242 GTGTTGTGGACTCTCTGTTGGAAAGTTTCATCTTGAACATGAGATGCTCAGGCC 301
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrrPleThrVal 120
Db 302 GACAAATCATAGAAATGGCTGGAAATTCATCCATCCACACCTCCAAATGAGATTAATCTGTA 361
QY 121 ProLeuThrIleAspThrTrrIleAlaValCysHisProLeuLeuYrHisThrValSer 140
Db 362 CCGTTAACATGACAGGTATATCACTGTCTGACCCGCTCAAGTACACACAGGTTCTCA 421
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 422 TACCCAGCCCGACCCGGAAAGTCATGTAGTGTTAATCATCACTCGCTTCGACACAGC 481
QY 161 IleProTyrTyrTrrPTrProAsnIleTrrPTrGluAspTyrIleSerThrSerValHis 180
Db 482 ATCCCTTATTAAGTGTGGCCCAACATCGGACTGAGAGACTACATCAGCATCTGTGCAT 541
QY 181 HisValLeuIleTrrPTrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 542 CACGTCTCATCTGAGATCCATGCTCTTACCGTCTTACCTGCTGCTTCATCTCTTC 601
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 602 ATCTTAATCATATCATTTGTGTACAGCTCAGAGAGAGAGACATTTTGCTCCCGTGGC 661
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 662 TACTCCACGGGGAAGACCAACCGCATCTGTTCACCATTAACCTCTTGTGCCACACTT 721
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 722 TGGGCCCCCGGACATCATAGATCTTACCACTTATAGGGGCGCCATCCAGAACCGC 781
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 782 TGGCTGTACATCATATGTCAGCATTCGCCACATCTGACCTTGTGAACACAGCATC 841
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 842 AACTCTCTCTCATCTTCACTCAGCAAGCCGTTCCGACCATGCGACCGCCAGCGCTC 901
QY 301 LysAlaPhePheLysGlyGlnLysGlnProValIleGlnPheTyrThrAsnHisAsnPheSer 320
Db 902 AAGGCTTTCTTCAGTCCAGAGAGCAACCTGTACAGTTCTACACCATATTAATCTTTCC 961
QY 321 IleThrSerSerProTrrPleSerProAlaSerSerHisCyBilelyMetLeuValTyr 340
Db 962 ATACAGAGAGAGCCCTGGATCTCGCCGCAACTCAGCATCAGAGATGGTGATC 1021
QY 341 GlnTyrAspLysAsnGlyLysProIleLeuValSerPro 353
Db 1022 CAGTATGACAAATAATGAGAAACCTATTAAGTATCCCCG 1060

RESULT 11
US-10-689-832-21
; Sequence 21, Application US/10689832
; Publication No. US20040121360A1
; GENERAL INFORMATION:
; APPLICANT: Majumder, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729D1V1
; CURRENT APPLICATION NUMBER: US/10/689,832
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835


```

/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: Patent Ver. 2.1
/ SEQ ID NO 21
/ LENGTH: 1343
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-689-832-21

```

```

Alignment Scores:
Pred. No.: 3,14e-177 Length: 1343
Score: 1868.00 Matches: 352
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 99.79% Indels: 0
DB: 17 Gaps: 0

```

US-09-995-225b-16 (1-353) x US-10-689-832-21 (1-1343)

```

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerLeuSerTrpSerProGly 20
DB 2 ATGGAGACACGACGACCCACCTGGACGACCAAGCTCGCTGTTGGTCCCGGCG 61
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuGlyGly 40
DB 62 TCGGCGTGGCGGCTTGGGTTTCCGCGCGGTCTACTACAGCTCTTGGTGGCTCGGT 121
QY 41 LeuProAlaAsnLeuLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGln 60
DB 122 TTACCGACAAATATCTTGACAGATCATCTCCACGCTGGTGGCAAGAGACAGAG 181
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPheIle 80
DB 182 TCCTCCTACACATATCTTGGACCTCGCTGCGGACATCTTGGCTCTTTTCATA 241
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 242 GGTGTGTGACCTTCTGTTGGAGATTTCATCTTGACATGACATGCTCAGGTCCC 301
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisGlySerIleTrpIleThrVal 120
DB 302 GACAAATCATGAAGTCTGTAATTTCTATTCATCCACCTCCATATGATTAATCTGTA 361
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 362 CGGTAAACCATTTGACAGGTATATCACTGTGCGCCCGGCTCAAGTACACAGCTCTCA 421
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 422 TACCAAGCCGCCACCCGAAAGTCAATGTATTAAGTATTAACATCTCTTCTGACAGC 481
QY 161 IleProTyrTyrTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
DB 482 ATCCCATATTAAGTGTGGCCCAATCATGTGACTAAACATCAACACCTCTGTGCAT 541
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200

```

```

DB 542 CAGCTCTATCTGGAACCACTGCTTACCGTCCCTGCTCCATCTTCTTC 601
QY 201 IleLeuAsnSerIleIleValTyrLysLeuAlaArgLysSerAsnPheArgLeuArgGly 220
DB 602 ATCTGAACATCATCTGTGTACAGCTCAGAGAGAAAGCAATTTTCGTCCGTGAGC 661
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 662 TACTCAGAGGAAAGCAACCGCATCTTGTTCACCATTAATCCATCTTGGCCACACTT 721
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
DB 722 TGGGCCCCCGCATCATATGATTTTACCACTCTATGGGGCCCGCATCCAGAACCGC 781
QY 261 TrpLeuValHisIleIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
DB 782 TGGCTGTATACATATATGTCGACATGCGCAACATGTAGCCCTTGGAAACAGGCATC 841
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
DB 842 AACTCTTCTCTACTGCTTCTATCAGCAGCGGTTCCGACCATGAGCGCCGACCGCTC 901
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
DB 902 AAGGCTTCTTCAAGTCCAGAACCAACTGTACAGTTCTACACCAATCATTAATCTTCC 961
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysGlyIleMetLeuValTyr 340
DB 962 ATTAACAAATTAACCCCTGATCTCGCGGCAAACTCAACATGCAAGATGCTGTGTAC 1021
QY 341 GlnTyrAspLysAsnGlyLysProIleLysValSerPro 353
DB 1022 CAGTATGCAAAATAATGAAAACTATTAAGTATTCGCCG 1060

RESULT 12
US-09-813-432-19
/ Sequence 19, Application US/09813432
/ Publication No. US20030148485A1
/ GENERAL INFORMATION:
/ APPLICANT: Taupier Jr., Raymond J
/ APPLICANT: Matmuder, Kamud
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Mezes, Peter S
/ APPLICANT: Verne, Corine A. M.
/ TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
/ FILE REFERENCE: 15966-729
/ CURRENT APPLICATION NUMBER: US/09/813,432
/ PRIOR FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 60/193,843
/ NUMBER OF SEQ ID NOS: 78

```



```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-813-432-19

Alignment Scores:
Pred. No.: 3,03e-176 Length: 1119
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.20% Indels: 0
DB: 10 Gaps: 0

US-09-995-225b-16 (1-353) x US-09-813-432-19 (1-1119)

QY 1 MetGluHfThRhiSAhIhIeUuAlaAlaAaSnSerLeuSerTrpTrpSerProGly 20
DB 1 ATGGAGCACACGACGCCCACTCGACGCAACAGCTCGCTGCTGGTGTGCTCCCGCC 60
QY 21 SerAlaCyGgIyLeuGlyPheValProValValTyTrpSerLeuLeuCySleuGly 40
DB 61 TCGGCGTGGCGGCTTGGGTTGCGGCCGCTGCTACTACAGCGCTTGTGCTGCTGCT 120
QY 41 LeuProAlaAaSnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgAlnLys 60
DB 121 TTACCAAGCAAAATCTTGACAGATGATCATCTCTCCAGCTGGTGGCAAGACAGAA 180
QY 61 SerSerTyraAntTyLeuLeuAlaLeuAlaAlaAaSnIleLeuValIleuPhePheIle 80
DB 181 TCTCTTCAACAATCTCTTGGCACTGCTGCTGCGCACTCTTGGCTCTTTTTCATA 240
QY 81 ValPheValAaPheLeuLeuGluAaPheIleLeuAaMetGlnMetProGlnValPro 100
DB 241 GTGTGTGTGACCTTCTGTTGGAAGATTTTCATCTTGAAACAGAGATGCTCAGGTC 300
QY 101 AaPlyAlleIleGluValLeuGluPheSerSerIleHfThRhiSerIleTrpIleThrVal 120
DB 301 GACAAATCATTAAGGCTGGAATTCATCATCATCACCTCCATATGATTAATCTGTA 360
QY 121 ProLeuThrIleAaPargTyIleAlaValCyAaHfProLeuTyTrpHfStrValSer 140
DB 361 CCGTTAACCTTAAGGATATGCTGCTGCTGCAACCTGCTCAAGTACCAAGGCTTCA 420
QY 141 TyProAlaArgThArgLysValIleValSerValTyTrpIleThrCySphLeuThrSer 160
DB 421 TACCAAGCCGCAACCGGAAGTCATGTAAAGTTTAACTACATCACTGCTTCTGAC 480
QY 161 IleProTyTrpTrpTrpProAaSnIleTrpThGluAaPlyTrpIleSerThrSerValHis 180
DB 481 ATCCCTTATTAATGCTGGCCCAACATCTGACATGAAACATCAACACCTCTGTGCAT 540
QY 181 HisValLeuIleTrpIleHisCySphThValTyTrpLeuValProCySerSerIlePhePhe 200
DB 541 CAGCTCTCATCTGGATCCACTGCTTCAACGCTTCACTGATGCTGCTGCTCACTTCTTC 600
QY 201 IleLeuAaSnSerIleIleValTyTrpLysLeuArgLysSerAaPheArgLysLeuArgLys 220
DB 601 ATCTTGAACATCAATGTTGTGTAACAAGCTCAGAGGAAAGACATTTTCTCTCCGTGC 660
QY 221 TySerThrGlySerThrThraAlaIleLeuPheThrIleHfSerIlePheAlaThrLeu 240
DB 661 TACTTCACGGGGAAGACCAACGCTTGTTCACCACTTCCATCTTGGCACACTT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTyTrpHfIleLeuTyTrpAlaProIleGlnAaArg 260
DB 721 TGGGCCCCCGCATCATGATCTTTTACCACTCTCATGGGGGCCCAACCAAGCCG 780
QY 261 TrpLeuValHisIleMetSerAaPlyAlaAaMetLeuAlaLeuLeuAaThraAlaIle 280
DB 781 TGGCTGTACATCATGTGTCGACATTCGCAACATGTCAACCTTCTGAAACAGACCATC 840

```

```

QY 281 AaPhePheLeuTyTrpCySphIleSerLysArgPheArgThrMetAlaAlaIleThrLeu 300
DB 841 AACTCTTCTCTTCAAGCTTCACTGATCAGCAAGCGGTTCGACCAATGCAAGCCGACGCTC 900
QY 301 LysAlaPhePheLysCySglnLysGlnProValGlnPheTyTrpAaSnHfAaPheSer 320
DB 901 AAGCTTTCTTCAAGGCGCAAGCAACCTGTACATTCATCAACCAATCACTTCTTCC 960
QY 321 IleThrSerSerProTrpIleSerProAlaAaSnHfCyAalleLysMetLeuValTy 340
DB 961 ATTAACAAGAGACCCCTGATCTCGCCGCAACTCACTGATCAAGAATGCTGTGTAC 1020
QY 341 GlnTyraPlyAaSnGlyLysProIleLys 350
DB 1021 CAGTATGACAAATATGAAAACCTATTAATA 1050

RESULT 13
US-10-174-364-19
; Sequence 19, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2
; CURRENT APPLICATION NUMBER: US/10/174,364
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1116)
US-10-174-364-19

Alignment Scores:
Pred. No.: 3,03e-176 Length: 1119
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.20% Indels: 0
DB: 15 Gaps: 0

US-09-995-225b-16 (1-353) x US-10-174-364-19 (1-1119)

QY 1 MetGluHfThRhiSAhIhIeUuAlaAlaAaSnSerLeuSerTrpTrpSerProGly 20
DB 1 ATGGAGCACACGACGCCCACTCGACGCAACAGCTCGCTGCTGGTGTGCTCCCGCC 60
QY 21 SerAlaCyGgIyLeuGlyPheValProValValTyTrpSerLeuLeuCySleuGly 40

```


QY 141 TyrProAlaArgThrArgIleValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCAGCCGGAAAGCATGTGTAAGTTTACATCACTGCTTCGACAGC 480
 QY 161 IleProTyrTyrTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTATTACTGGTGGGCCCAACATCTGAGCTGAAAGCTACATCAGACCTCTGTGCAT 540
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CAGCTCCTCATCTGGATTCACCTGCTTACCGCTGAGTGGCCGCTCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrIleValArgIleCysSerAsnPheArgLeuArgIle 220
 DB 601 ATCTTAACATCAATCATTTGTGTCAAGCTCAGAGGAAAGCAATTTCTCTCCGTCGCG 660
 QY 221 TyrSerThrGlyIleValThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCCAGGGGAAAGCAACCCGCACTTCTTCACTTACCTTCATCTTGGCACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValaProIleGlnAsnArg 260
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTTACACCTGATGGGGGCCCATCCAGAACCGC 780
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
 DB 781 TGGCTGTGTACATCATGTGTCCGACATTCGCAACATGTGTACCTTCTGAAACAGCCATC 840
 QY 281 AsnPhePheLeuTyrCysPheIleSerIleValArgPheArgThrMetAlaAlaThrLeu 300
 DB 841 AACTCTCTCTCACTGCTCATCATCAGAACGGGTTCGCAACATGGAGCCGCGCACGCTC 900
 QY 301 LysAlaPhePheLeuCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
 DB 901 AAGGCTTTCTTCAAGTGCAGAAAGCAACCTGTACAGTTTCAACCATCATTAATCTTTCC 960
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLeuMetLeuValTyr 340
 DB 961 ATTAACAAGTAGCCCTGGATCTCGCGGCAACTCATCATCAAGATGCTGTGTAC 1020
 QY 341 GlnTyrAspLysAsnGlyLysProIleLys 350
 DB 1021 CAGTATGACAAATAAGAAACCTTAATAAA 1050
 RESULT 15
 US-10-689-832-19
 ; Sequence 19, Application US/10689832
 ; Publication No. US20040121380A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Majumder, Kamud
 ; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
 ; FILE REFERENCE: 15966-729DI1
 ; CURRENT APPLICATION NUMBER: US/10/689,832
 ; PRIORITY FILING DATE: 2003-10-20
 ; PRIOR APPLICATION NUMBER: 09/813,432
 ; PRIOR FILING DATE: 2001-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,835
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,768
 ; PRIOR FILING DATE: 2000-03-20
 ; PRIOR APPLICATION NUMBER: 60/190,972
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,199
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/191,947
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: 60/192,665
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,657
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: 60/192,984
 ; PRIOR FILING DATE: 2000-03-28

; PRIOR APPLICATION NUMBER: 60/192,664
 ; PRIOR FILING DATE: 2000-03-28
 ; Remaining prior Application data removed - See file Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 19
 ; LENGTH: 1119
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-689-832-19
 Alignment Scores:
 Pred. No.: 3,03e-176 Length: 1119
 Score: 1857.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.20% Indels: 0
 DB: 17 Gaps: 0
 US-09-995-225b-16 (1-353) x US-10-689-832-19 (1-1119)
 QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
 DB 1 ATGGAGCACACGCGACGCCACCTCGCAGCCAAACAGCTCGCTGTGGTGTCCCGCGC 60
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly 40
 DB 61 TCGGCTGGGGCTGGTGGTTCGGCCGCGGTGTCTACAGCTCTTGTGCTGCTCGGT 120
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgLys 60
 DB 121 TTACGCAAAATATCTTGACAGTATCATCTTCCAGCTGTGGGCAAGAAACAGAG 180
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
 DB 181 TCCTCTACACATCATCTTGGCATCGCTGCGCGGAGCATCTGTGCTCTTTTCATA 240
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
 DB 241 GTGTGTGTGACTTCTGTGTGAAGATTCATCTTGAAGATGAGAGTCCAGTCCCGC 300
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
 DB 301 GACAAATCATAGAAATGCTGGAAATTCATTCATCATCACCTCCATATGATTAATCTGTA 360
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
 DB 361 CGCTTAACATTTGACAGGTATATCGCTGTCTGCACCGCTCAAGTACACAGGTCTCA 420
 QY 141 TyrProAlaArgThrArgIleValIleValSerValTyrIleThrCysPheLeuThrSer 160
 DB 421 TACCCAGCCCGCAGCCGGAAAGCATGTGTAAGTTTACATCACTGCTTCGACAGC 480
 QY 161 IleProTyrTyrTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
 DB 481 ATCCCTATTACTGGTGGGCCCAACATCTGAGCTGAAAGCTACATCAGACCTCTGTGCAT 540
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
 DB 541 CAGCTCCTCATCTGGATTCACCTGCTTACCGCTGAGTGGCCGCTCATCTTCTTC 600
 QY 201 IleLeuAsnSerIleIleValTyrIleValArgIleCysSerAsnPheArgLeuArgIle 220
 DB 601 ATCTTAACATCAATCATTTGTGTCAAGCTCAGAGGAAAGCAATTTCTCTCCGTCGCG 660
 QY 221 TyrSerThrGlyIleValThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
 DB 661 TACTCCAGGGGAAAGCAACCCGCACTTCTTCACTTACCTTCATCTTGGCACACTT 720
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValaProIleGlnAsnArg 260
 DB 721 TGGGCCCCCGCATCATCATGATTTCTTTACACCTGATGGGGGCCCATCCAGAACCGC 780

```
Oy 261 TrrleuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
    |||
Db 781 TGCTGGTACACATCATGTGCCACATGTGCAGCCCTTGTGAACACAGCCATC 840
Oy 281 AsnPhenLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaIleThrLeu 300
    |||
Db 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGCACCATGGCAGCCGCGCTC 900
Oy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheser 320
    |||
Db 901 AAGGCTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATTAATTTTCC 960
Oy 321 IleThrSerSerProTrrPileSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
    |||
Db 961 ATACCAAGTAGCCCTGGATCTCCGCCGCAAACTCAACCTGCATCAAGATGCTGGTGTAC 1020
Oy 341 GlnTyrAspLysAsnGlyLysProIleLys 350
    |||
Db 1021 CAGTATGACAAATAATGAAAAAAGCTATATAAA 1050
```

Search completed: December 14, 2004, 22:25:44
Job time : 465 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:04:39 ; Search time 15.4373 Seconds

(Without alignments)
2200.157 Million cell updates/sec

Title: US-09-995-225B-16

Perfect score: 1872
Sequence: 1 MEHTAHILAAANSLSLWMSPG.....CIKMLVYQDKKPKIKVSP 353Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR 79: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	242	12.9	411	2	156444
2	242	12.9	412	2	S23436
3	241	12.9	393	2	A39251
4	238.5	12.7	398	2	JN0708
5	231.5	12.4	355	2	A45177
6	226	12.1	352	2	JE0296
7	225	12.0	357	2	UC7319
8	219.5	11.7	388	2	JN0605
9	219	11.7	384	2	A47249
10	218	11.6	423	2	UC7677
11	217.5	11.6	340	2	TI8704
12	217.5	11.6	394	2	UC7209
13	217.5	11.6	658	2	JC8011
14	217	11.6	367	2	UC2421
15	217	11.6	367	2	J49022
16	217	11.6	367	2	I56520
17	214.5	11.5	595	2	JC8012
18	213.5	11.4	355	2	I49339
19	213	11.4	352	2	A45747
20	213	11.4	370	2	S43087
21	213	11.4	384	2	JC4629
22	212	11.3	360	2	T23063
23	211.5	11.3	376	2	T24368
24	211.5	11.3	380	2	UC2338
25	211.5	11.3	504	2	T29338
26	210.5	11.2	380	2	S36143
27	209	11.2	424	2	JH0164
28	208.5	11.1	363	2	I57940
29	207.5	11.1	399	2	A46632

30	207.5	11.1	428	2	JN0692	cholecystokinin ty
31	206.5	11.0	380	2	A48227	kappa opioid recep
32	206.5	11.0	380	2	UC2434	kappa opioid recep
33	206.5	11.0	391	2	A41795	somatostatin recep
34	206.5	11.0	391	2	C41795	somatostatin recep
35	205.5	11.0	353	2	JC2492	G protein-coupled
36	205.5	11.0	391	2	A39297	somatostatin recep
37	205	11.0	355	2	G02436	chemokine (C-C) re
38	205	11.0	355	2	UC5067	G protein-coupled
39	205	11.0	418	2	S59601	vasopressin recep
40	204.5	10.9	352	2	S60024	bradykinin Bi rece
41	204.5	10.9	380	2	A55259	kappa opioid recep
42	204	10.9	353	2	S28787	neuropeptide Y/pep
43	204	10.9	372	2	I38532	delta opioid recep
44	203.5	10.9	314	2	TI8831	hypothetical prote
45	203.5	10.9	477	2	UC7913	capa receptor (CGI

ALIGNMENTS

RESULT 1

156444
thyrotrophin-releasing hormone receptor - mouse
C:Species: Mus sp. (mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 11-Jan-2000
C:Accession: I56444
R:Seller, R.E.; Taylor, P.L.; Lamb, R.F.; Zabaynik, J.; Anderson, L.; Eldre, K.A.
J. Mol. Endocrinol. 10, 199-206, 1993
A:Title: Functional expression and molecular characterization of the thyrotrophin-relea
A:Reference number: I56444, PMID:93249585, PMID:8387312
A:Accession: I56444
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-411 <RES>
A:Cross-references: GB:S60053; NID:G300151; PIDN:AB26491.1; PID:G300152
C:Superfamily: adenosine receptor A1

Query Match

Best Local Similarity 12.9%; Score 242; DB 2; Length 411;
Matches 83; Conservative 71; Mismatches 125; Indels 88; Gaps 14;

QY	19	PGSACGLGF-VPVVYVSLLC-IGLPANILTYITISQVARRQKSSYN-VYLALAAADIL 75
DB	16	POVAVALLEYOVVITLLVAVICGIVGIMVAVLVMR--TKGMRITNCLVSLAVADLM 73
QY	76	VLPEFIVFVDFLEDFITLNMQMPQVDPKI-----TEVLEFSSIHSTIWTVP 122
DB	74	VL-----VAAGLRITDSIYGSWYGVGCLCTIYQYLGIMASSCITAF 119
QY	123	TIDRYIAVCHPLKHTVSYPARTKVIYSVYITCFITSIPIYVW-----PNIWT-BDYIS 176
DB	120	TIERYIAICHPKIAQCLCFPSRAKIIIFVMA---FTSIYCMIMFLDLNISTYDAIV 176
QY	177	TSVHNLVIMHCTVYLVVCSIEFIIINSIIVKL----- 210
DB	177	ISCGYKISRNYSPIITLMDGVFVPMILATVLVGFARILFLANDIPSDPKENSTWKN 236
QY	211	---RRKSNRLRG-----YSTGKTAIFITISIPATMAP-RIMILVHLGAPI 257
DB	237	DSTHQKNNMLNLTTRNCRPSYSSRKQVTKMLAVVILFALLMPPRTLVVANSFSSPF 296
QY	258	QNRMLVHINSDIANMLALNTAINFELYCFISKRFTMAATLKAEFKQKQVQPYTNH 317
DB	297	QENWFLL---FCRICIYNSAINPIVYIMSQKFR---AARFKLCNCKQKPTERANV 348
QY	318	NFSITSS 324
DB	349	SVALNTS 355

RESULT 2

S23436

thyrotropin receptor - rat
 N/Alternate names: thyrotropin-releasing hormone receptor
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C/Accession: S23436; I53279; A49168; P00326
 R/de la Pena, P.; Delgado, L.M.; del Camino, D.; Barrios, F.
 Biochem. J. 284, 891-899, 1992
 A/Title: Cloning and expression of the thyrotropin-releasing hormone receptor from GH(3)
 A/Reference number: S23436; MUID:92322017; PMID:1377915
 A/Accession: S23436
 A/Molecule type: mRNA
 A/Residues: 1-412 <PEN>
 A/Cross-references: UNIPROT:Q01717; EMBL:X64630; NID:957394; PDB:CA45913.1; PID:957395
 R/Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N.
 Endocrinology 134, 432-440, 1994
 A/Title: Batrachoiodin transcriptionally and posttranscriptionally up-regulates thyrotropin-
 A/Reference number: I53279; MUID:94102223; PMID:8215956
 A/Accession: I53279
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-412 <RES>
 A/Cross-references: GB:D17469; NID:9464199; PDB:BAA0289.1; PID:9464200
 R/Zhao, D.; Yang, J.; Jones, K.E.; Gerald, C.; Suzuki, Y.; Hogan, P.G.; Chin, W.W.; Tash
 Endocrinology 130, 3529-3536, 1992
 A/Title: Molecular cloning of a complementary deoxyribonucleic acid encoding the thyrotro
 A/Reference number: A49168; MUID:92283212; PMID:1317787
 A/Accession: A49168
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-12, 'D', 14-290, 'K', 292-412 <ZHA>
 A/Experimental source: GH cells
 A/Note: Sequence extracted from NCBI backbone (NCBIN:104788, NCBI:P104795)
 R/Yamada, M.; Monden, T.; Satch, T.; Iizuka, M.; Murakami, M.; Iriuchijima, T.; Mori, M.
 Biochem. Biophys. Res. Commun. 184, 367-372, 1992
 A/Title: Differential regulation of thyrotropin-releasing hormone receptor mRNA levels b
 A/Reference number: P00326; MUID:92231953; PMID:1373613
 A/Accession: P00326
 A/Molecule type: mRNA
 A/Residues: 30-58, 'P', 60-222, 'T', 224-261 <YAM>
 A/Experimental source: strain Wistar
 A/Note: the authors translated the codon ACA for residue 88 as Ala
 C/Superfamily: adenosine receptor A1
 C/Keywords: G protein-coupled receptor; transmembrane protein
 F/1-22/Domain: transmembrane #status predicted <TM1>
 F/32-94/Domain: transmembrane #status predicted <TM2>
 F/71-92/Domain: transmembrane #status predicted <TM3>
 F/116-140/Domain: transmembrane #status predicted <TM4>
 F/165-186/Domain: transmembrane #status predicted <TM5>

Query Match 12.9%; Score 242; DB 2; Length 412;
 Best Local Similarity 22.6%; Pred. No. 1.1e-12;
 Matches 83; Conservative 71; Mismatches 125; Indels 88; Gaps 14;

QY 19 PSAGCIGF-VPVYVYSLLC-IGLPANILTYIISQVARRQKSSYN-YLALAAADIL 75
 DB 16 PVAVALRYQVVTILVITCGIGVGNIMVAVLVMR--TKMRATPCYVLSLAVADLM 73
 QY 76 VLFVIFVDFLEDFILNMQMPQVDPKI-----IEVLEFSSIHSTSIWITVPL 122
 DB 74 VL-----VAAGLPNITDSIYGSWVGYGCLCTITLYQYGINASSCSITAF 119
 QY 123 TIDRIYAVCHPLKHTVTSYPAKTRKYVSVYITCFLTSIPYWM-----PNIMT-EDYIS 176
 DB 120 TIERIYAIACHPIKAQFLCTFSRAKIIIFVWA---FTSIYCMLEPFLDLNISTYKDAIV 176
 QY 177 TSVHVLWIMHCFYVIVPCISFFILNSIYVKL----- 210
 DB 177 ISCGKISRNTYSPYILMDGVFVPMILATVIGFIARILFLNPISDPKENSCKTKN 236
 QY 211 ---RRKSNFRLG-----YSTGKTTALFTITSIFATLMAP-RITMILYHLYGAPI 257
 DB 237 DSTHQNKMNLNATNRCFNSITVSRKQYTKMLAVVVIIFALLMIMYKRLVAVNSLSSPF 296

QY 258 QNRVLVHMSDIANMLALNTAINFPLCYFSKRRPTMAATLKAFFKQCKQPOVFTYTH 317
 DB 297 QENWFLL-----FCRCIYINSAINPVIYVIMSQKRR-----AAFRKLCNCKQKPTKKAANY 348
 QY 318 NFSITSS 324
 DB 349 SVALNYS 355

RESULT 3
 A39251
 thyrotropin-releasing hormone receptor - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
 C/Accession: A39251
 R/Staub, R.E.; Frech, G.C.; Joho, R.H.; Gershengorn, M.C.
 Proc. Natl. Acad. Sci. U.S.A. 87, 9514-9518, 1990
 A/Title: Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releasin
 A/Reference number: A39251; MUID:91088548; PMID:2175902
 A/Accession: A39251
 A/Molecule type: mRNA
 A/Residues: 1-393 <STR>
 A/Cross-references: UNIPROT:P21761; GB:M59811; GB:M37490; NID:9202153; PDB:AAA40480.1;
 C/Superfamily: adenosine receptor A1
 C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 12.9%; Score 241; DB 2; Length 393;
 Best Local Similarity 22.1%; Pred. No. 1.3e-12;
 Matches 81; Conservative 75; Mismatches 123; Indels 88; Gaps 15;

QY 19 PSAGCIGF-VPVYVYSLLC-IGLPANILTYIISQVARRQKSSYN-YLALAAADIL 75
 DB 16 PVAVALRYQVVTILVITCGIGVGNIMVAVLVMR--TKMRPTPCYVLSLAVADLM 73
 QY 76 VLFVIFVDFLEDFILNMQMPQVDPKI-----IEVLEFSSIHSTSIWITVPL 122
 DB 74 VL-----VAAGLPNITDSIYGSWVGYGCLCTITLYQYGINASSCSITAF 119
 QY 123 TIDRIYAVCHPLKHTVTSYPAKTRKYVSVYITCFLTSIPYWM-----PNIMT-EDYIS 176
 DB 120 TIERIYAIACHPIKAQFLCTFSRAKIIIFVWA---FTSIYCMLEPFLDLNISTYKNAV 176
 QY 177 TSVHVLWIMHCFYVIVPCISFFILNSI---IYV----- 208
 DB 177 VSCGKISRNTYSPYILMDGVFVPMILATVIGFIARILFLNPISDPKENSCKTKN 236
 QY 209 -KLRRKSNFRLG-----YSTGKTTALFTITSIFATLMAP-RITMILYHLYGAPI 257
 DB 237 DSIHQNKMNLNATNRCFNSITVSRKQYTKMLAVVVIIFALLMIMYKRLVAVNSFLSSP 296
 QY 258 QNRVLVHMSDIANMLALNTAINFPLCYFSKRRPTMAATLKAFFKQCKQPOVFTYTH 317
 DB 297 QENWFLL-----FCRCIYINSAINPVIYVIMSQKRR-----AAFRKLCNCKQKPTKKAANY 348

QY 318 NFSITSS 324
 DB 349 SVALNYS 355

RESULT 4
 JN0708
 thyrotropin-releasing hormone receptor - human
 C/Species: Homo sapiens (man)
 C/Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
 C/Accession: S40682; JN0759; S50151; S50152; I38356; JN0708
 R/Matire, V.; Karlsson, H.B.; Wright, M.S.; Lundell, I.; Fjeldheim, A.K.; Gabrielsen, O.S.
 Biochem. Biophys. Res. Commun. 195, 179-185, 1993
 A/Title: Molecular cloning of a functional human thyrotropin-releasing hormone receptor.
 A/Reference number: S40682; MUID:93371401; PMID:8395824
 A/Accession: S40682
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-398 <MAT>

A:Cross-references: UNIPROT:P31981; EMBL:X75071; NID:G404157; PIDN:CAA52965.1; PID:G404157
R.Yamada, M.; Monden, T.; Satoh, T.; Satoh, N.; Murakami, M.; Iriuchijima, T.; Kakegawa, S.
Biochem. Biophys. Res. Commun. 193, 737-745, 1993
A>Title: Pituitary adenomas of patients with acromegaly express thyrotropin-releasing hormone receptor gene.
A:Reference number: JN0759; MUID:93384596; PMID:8396925
A:Accession: JN0759
A:Molecule type: mRNA
A:Residues: 1-398 <YMA>
A:Cross-references: GB:016945; NID:G577631; PIDN:BA04120.1; PID:G577632
R.Hinuma, S.; Hosoya, M.; Ogi, K.; Tanaka, H.; Nagai, Y.; Onda, H.
Biochim. Biophys. Acta 1219, 251-259, 1994
A>Title: Molecular cloning and functional expression of a human thyrotropin-releasing hormone receptor.
A:Reference number: S50151; MUID:95002135; PMID:7918619
A:Accession: S50151
A:Molecule type: preliminary
A:Status: preliminary
A:Residues: 1-263 <HIN>
A:Accession: S50152
A:Molecule type: preliminary
A:Status: preliminary
A:Residues: 267-398 <HI2>
R.Duchie, S.M.; Taylor, P.L.; Anderson, L.; Cook, J.; Bidne, K.A.
Mol. Cell. Endocrinol. 95, R11-R15, 1993
A>Title: Cloning and functional characterisation of the human TRH receptor.
A:Reference number: I38356; MUID:94063224; PMID:8243797
A:Accession: I38356
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-398 <RES>
A:Cross-references: EMBL:X72089; NID:G440155; PIDN:CAA50979.1; PID:G440156
C:Genetics:
A:Gene: GDB:TRHR
A:Cross-references: GDB:228955; OMIM:188545
A:Map position: 8q23-8q23
C:Superfamily: adenosine receptor A1
C:Keywords: G protein-coupled receptor; receptor; transmembrane protein
F:29-51/Domain: transmembrane #status predicted <TM1>
F:62-83/Domain: transmembrane #status predicted <TM2>
F:101-121/Domain: transmembrane #status predicted <TM3>
F:146-166/Domain: transmembrane #status predicted <TM4>
F:194-215/Domain: transmembrane #status predicted <TM5>
F:267-288/Domain: transmembrane #status predicted <TM6>
F:297-319/Domain: transmembrane #status predicted <TM7>

Query Match	12.7%	Score 238.5;	DB 2;	Length 398;
Best Local Similarity	22.4%	Pred. No. 2.1e-12;		
Matches 80;	Conservative 70;	Mismatches 120;	Indels 87;	Gaps 13;
28	VPVVVYSLLC-IGLPANILTVIIISQLVAROKSSV-YLLAALADIVLFFIVDF	85		
DB	VVTILVLLVLIIGLGIVGNIVLVLMR--TKMPTNVCYVSLVADLMV-----	75		
86	LLEDPIINQMPOVPDKI-----IEVLEFSIHITSITWITVPLTIDRYIAVCH	132		
DB	-----VAAGLPNITDSIVGSWVYGVCCLITLQYIGINASSCSITAFTERIATICH	129		
133	PLKHTVTSVPARFRKRVIVSYITCFLTSIPYTW-----PINWT-EDYISVSHVHLWI	186		
DB	130 PIKAFPLCTFSRAKKIIIFVWA--FTSLXCMLEFLLDNIISTYKDAIVISCGYKISHN	186		
187	HCFTVLVCSIFFIINSIIVYKL-----RRKSN-----	215		
DB	187 YISPIYLMFGVYVYVPMILATVLYGFIARLLFLNPIPSDKENSKTWKNDSTHOVNLN	246		
216	-----FLRGYSTKTAILEFTITSIPATLMAP-RIIMILYMLYGAPIQRMVLVHMS	267		
DB	247 VNTSNRCFNSVTSRRQVTKGLAVVILFALLMMPYRFLVVVNSFLSPQGNWFL----	303		
266	DIANMLALNTAINFLFYCISKRFRITMAATLKAFFCCQKQOPVPTYNHNFSTSS	324		
DB	304 -FCRIYIYNSAINPVYINLMSOKF-----AARFLCNCKCKPPEKPNVYSVALNYS	355		

RESULT 5

A45177

Chemokine (C-C) receptor 1 - human

N|Alternate names: C-C CKR-1; macrophage inflammatory protein-1-alpha receptor

C|Species: Homo sapiens (man)

C|Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C|Accession: A45177; 155671

R|Nucleotide, K. J. Digregorio, D. J. Mak, J. Y. Horuk, R. J. Schall, T. J. Cell 72, 415-425, 1993

A|Title: Molecular cloning, functional expression, and signaling characteristics of a C

A|Reference number: A45177; MUID:93161416; PMID:7679328

A|Accession: A45177

A|Status: nucleic acid sequence not shown

A|Molecule type: mRNA

A|Residues: 1-355 <NEO>

A|Cross-references: UNIPROT:P32246; GB:L10918; NID:g292416; PIDN:AAA36543.1; PID:g29241

A|Experimental source: HL60 cells

A|Note: sequence extracted from NCBI backbone (NCBIP:124876)

R|Gao, J.

J. Exp. Med. 177, 1421-1427, 1993

A|Title: Structure and functional expression of the human macrophage inflammatory 1 alp

A|Reference number: 155671; MUID:93240122; PMID:7683036

A|Accession: 155671

A|Status: preliminary; translated from GB/EMBL/DBJ

A|Molecule type: mRNA

A|Residues: 1-355 <RES>

A|Cross-references: GB:L10918; NID:g292416; PIDN:AAA36543.1; PID:g292417

C|Genetics:

A|Gene: GDB:CMKR1; CMKR-1

A|Cross-references: GDB:138446; OMIM:601159

A|Map position: 3p21-3p21

C|Superfamily: vertebrate rhodopsin

C|Keywords: disulfide bond; G protein-coupled receptor; glycoprotein; phosphoprotein; t

F|36-60/Domain: transmembrane #status predicted <TM1>

F|71-91/Domain: transmembrane #status predicted <TM2>

F|108-129/Domain: transmembrane #status predicted <TM3>

F|147-171/Domain: transmembrane #status predicted <TM4>

F|205-223/Domain: transmembrane #status predicted <TM5>

F|240-264/Domain: transmembrane #status predicted <TM6>

F|288-305/Domain: transmembrane #status predicted <TM7>

F|5/Binding site: carbohydrate (Asn) (covalent) #status predicted

F|24-273,106-183/disulfide bonds: #status predicted

F|345/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted

Query Match 12.4%; Score 231.5; DB 2; Length 355;

Best Local Similarity 26.3%; Pred. No. 7.3e-12;

Matches 81; Conservative 57; Mismatches 105; Indels 65; Gaps 13;

QY 22 AGLGFPVVTYYSLTCLGEPANITVYIISQVARRQKSYNYTLAAADIVLPLP 81

DB 31 AFGQLLPPL-YSLVPIYIGVGNILVVLVQYRKLKMTSI-VLMLAISDLLFLPLP 88

QY 82 F-VDF-LLEDFILMMQPOVDKRIEVLFPSSHTSIWITVPLTIDRYIAYCH--PLKY 136

DB 89 FWDYKAKDWWPBGDA-----CKILSGFYTGYSSEFFIILLITDRIAYLVHAVPALRA 144

QY 137 HTVSYPARTKRVISVYITGELTSP-YVWMPNITWEDYISTSVH----- 181

DB 145 RTVTFPGVITSLIIMALAI---LASMPLGYFSKTQWEPHTNHTCSLHPFHSREWKLFQAL 201

QY 182 -----VLMIHCFYTVLVGCSIFLINSIIVYKLRKSKSFLRLGISTGTMLLF 231

DB 202 KLVPLGLVPLVPLVITCYT-----GIKKILRRPNE-----KKSAVRILF 242

QY 232 TITSIPATLWAPRIMITLHYGAPI-----QNRMLVHIMSIDANMLALINTAINFLY 285

DB 243 VIMIIIFLFWPFPYVLTILISVQDFLFTHEBQGRHL-DLAVQYTVIAYTHCCVNPVIY 301

QY 286 CFISKRR 293

DB 302 AFVGRFR 309

RESULT 6

JE0296

Chytrotropin releasing hormone receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004

C/Accession: JE0296

R:Radant, H.; Nakamura, T.; Itoh, J.; Iwasa, H.; Kanatani, A.; Borkowski, J.; Ihara, Y.

Biochem. Biophys. Res. Commun. 250, 66-71, 1998

A>Title: Cloning and characterization of a new subtype of chytrotropin-releasing hormone

A:Reference number: JE0296; MUID:98407892; PMID:9735333

A:Accession: JE0296

A:Molecule type: mRNA

A:Residues: 1-352 <ITA>

A:Cross-references: UNIPROT:O68820; DDBJ:AB015645; NID:93660553; PIDD:BA33437.1; PIDD:93

C:Superfamily: adenosine receptor A1

F:26-46/Domain: transmembrane #status predicted <TM1>

F:58-80/Domain: transmembrane #status predicted <TM2>

F:97-118/Domain: transmembrane #status predicted <TM3>

F:142-165/Domain: transmembrane #status predicted <TM4>

F:188-209/Domain: transmembrane #status predicted <TM5>

F:252-273/Domain: transmembrane #status predicted <TM6>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

F:282-304/Domain: transmembrane #status predicted <TM7>

C:Superfamily: endothelin receptor B

C:Keywords: extracellular protein; glycoprotein; neuropeptide; transmembrane protein

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

Query Match 12.0%; Score 225; DB 2; Length 357;

Best Local Similarity 24.9%; Pred. No. 2, 5e-11;

Matches 87; Conservative 70; Mismatches 112; Indels 80; Gaps 18;

QY 245 IIMILYHLYGAPIONRKL-----VHMSDIANMLALNTAINFLYCFISKRRPTMA 236
 DB 340 HAQRLAAVGLININIGISDAFNDYFRIDYTSVLYFLSTCINPLNIMSHKREAR 399
 QY 297 AATLKAFKCKQK----QPVQFTYNNHPS 320
 DB 400 KITLTROFGIARNNHHQSQSH-QHNYS 426

RESULT 14

JC2421
 opioïd receptor homolog, MOR-C - mouse

C/Species: Mus musculus (house mouse)
 C/Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
 C/Accession: JC2421; 149122
 R/1stnl, M.; Takeshima, H.; Mori, M.; Nakagawara, K.; Takeuchi, T.
 Biochem. Biophys. Res. Commun. 205, 1353-1357, 1994
 A/Title: Structure and chromosomal mapping of gene for the mouse kappa-opioid receptor
 A/Reference number: JC2434; MUID:95100967; PMID:7802669
 A/Accession: JC2421
 A/Molecule type: mRNA
 A/Residues: 1-367 <NIS>
 A/Cross-references: UNIPROT:P35377; DDBJ:D31663
 R/Halford, W.P.; Gebhardt, B.M.; Carr, D.J.
 J. Neuroimmunol. 59, 91-101, 1995
 A/Title: Functional role and sequence analysis of a lymphocyte orphan opioïd receptor.
 A/Reference number: 149122; MUID:95318231; PMID:7797625
 A/Accession: 149122
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: mRNA
 A/Residues: 1-357 <RES>
 A/Cross-references: EMBL:U14165; NID:9540092; PIDN:AAA87899.1; PID:9540093
 A/Genetic: MOR-C; OOR
 A/Map position: 2H2-4
 A/Introns: 75/2; 194/1
 C/Superfamily: vertebrate rhodopsin
 C/Keyword: receptor

Query Match 11.6%; Score 217; DB 2; Length 367;
 Best Local Similarity 22.8%; Pred. No. 1.2e-10;
 Matches 74; Conservative 65; Mismatches 117; Indels 68; Gaps 13;
 QY 7 HLLAANGSLGWSGSAAGLGF--VPVYYSLLLCL-GLPANILTVIISQIVARROKSSYN 64
 DB 32 HLLANSHSAFLP-----LGLKTVIGLYLAVCIGLGNCLVMYVI--LRHTKMTATN 84
 QY 65 -YLLAALAADILVLFPIVF--VDFLEDFILNMQMPQVPDKIEVLEFSSIHSTIWTVP 121
 DB 85 IYFNLALADTLVLTLPFGSTDIILGFMPFGNAL-----CKTVIADYMMFTSTFTLTA 140
 QY 122 LITDRYIANCHPKYHTVSPARTRKIVYSVITCTLSIPY----- 163
 DB 141 MSVDRYVAICHPIRALDVRTSSKAQAVNAIMALASVGVPAIMGSAQVEDEIECTIVE 200
 QY 164 -----YMWPNITWEDYISTSVHVLIMHCFVYVLPSCIFPLINSIIYKLR----- 211
 DB 201 IPAPQDYWGP-----VPAICIFLPSF--IIPVLIISVCYSLMIRRLRGVRLLS 246
 QY 212 --RKSNFLRGYSTGKTTAILFTTISIFATLMAPRIIMILYHLYGAPIONRMLVHIMSDI 269
 DB 247 GSREKDRNLR-----RITRLVLVVAVFVGCMTPVQVFLVQGLGVQPGSETAVAILR-F 300
 QY 270 ANMLALNTAINFLYCFISKRRP 293
 DB 301 CTALGYVNSCLNPIILYAFIDENFK 324

RESULT 15
 149022
 kappa opioïd receptor 3 - mouse
 C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C/Accession: 149022
 R/Pan, Y.X.; Cheng, J.; Xu, J.; Rossi, G.; Jacobson, E.; Ryan-Moro, J.; Brooks, A.I.; D
 Mol. Pharmacol. 47, 1180-1188, 1995
 A/Title: Cloning and functional characterization through antisense mapping of a kappa 3
 A/Reference number: 149022; MUID:95327076; PMID:7603458
 A/Accession: 149022
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: mRNA
 A/Residues: 1-367 <RES>
 A/Cross-references: UNIPROT:P35377; EMBL:U09421; NID:9551484; PIDN:AAA81333.1; PID:9551
 C/Superfamily: vertebrate rhodopsin

Query Match 11.6%; Score 217; DB 2; Length 367;
 Best Local Similarity 22.8%; Pred. No. 1.2e-10;
 Matches 74; Conservative 65; Mismatches 117; Indels 68; Gaps 13;
 QY 7 HLLAANGSLGWSGSAAGLGF--VPVYYSLLLCL-GLPANILTVIISQIVARROKSSYN 64
 DB 32 HLLANSHSAFLP-----LGLKTVIGLYLAVCIGLGNCLVMYVI--LRHTKMTATN 84
 QY 65 -YLLAALAADILVLFPIVF--VDFLEDFILNMQMPQVPDKIEVLEFSSIHSTIWTVP 121
 DB 85 IYFNLALADTLVLTLPFGSTDIILGFMPFGNAL-----CKTVIADYMMFTSTFTLTA 140
 QY 122 LITDRYIANCHPKYHTVSPARTRKIVYSVITCTLSIPY----- 163
 DB 141 MSVDRYVAICHPIRALDVRTSSKAQAVNAIMALASVGVPAIMGSAQVEDEIECTIVE 200
 QY 164 -----YMWPNITWEDYISTSVHVLIMHCFVYVLPSCIFPLINSIIYKLR----- 211
 DB 201 IPAPQDYWGP-----VPAICIFLPSF--IIPVLIISVCYSLMIRRLRGVRLLS 246
 QY 212 --RKSNFLRGYSTGKTTAILFTTISIFATLMAPRIIMILYHLYGAPIONRMLVHIMSDI 269
 DB 247 GSREKDRNLR-----RITRLVLVVAVFVGCMTPVQVFLVQGLGVQPGSETAVAILR-F 300
 QY 270 ANMLALNTAINFLYCFISKRRP 293
 DB 301 CTALGYVNSCLNPIILYAFIDENFK 324

Search completed: December 14, 2004, 20:25:45
 Job time : 16.4373 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 14, 2004, 20:28:19 ; Search time 2566 Seconds
(without alignment) 4974.178 Million cell updates/sec

Title: US-09-995-225b-16

Perfect score: 1872
Sequence: 1 MHTAHILANSLSLWSMSPG.....CIKMLVYDYDKNGKPIKVP 353

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+pn.model -DEV=xlh
-Q/cg2.1/USPTO.spoc1/US0999225/runat.14122004.110733.18040/app.query.fasta_1.519
-DB=EST -QFMT=fastlap -SUFFIX=est -MIMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0999225@cgn1.1.2607@runat.14122004.110733.18040 -NCPU=6 -ICPU=3
-NO WMAP -LARGEDUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST :
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	925	49.4	803	9	CC905917 t025011ba
2	606	32.4	806	9	CNS02284 AL178069 Tetracodon
3	509	27.2	824	9	CL137737 ISB1-110A
4	506	27.0	362	8	AZ513631 IM0359L07
5	453.5	24.2	947	9	CNS02BVO AL190367 Tetracodon
6	444	23.7	867	9	CL137848 ISB1-110C
7	433.5	23.2	813	8	BZ169352 CH230-259
8	426	22.8	898	9	CNS02ENP AL193966 Tetracodon
9	414.5	22.1	515	1	AL921815 AL921815

C 10	412	22.0	448	8	AQ225693	AQ225693 HS 2009 B
11	380	20.3	1269	5	CL649540	CH213-237
12	359	19.2	708	5	BUS10327	BUS10327 UI-M-DJ2-
C 13	314	16.8	588	8	AQ344035	AQ344035 RBC11-12
14	309	16.5	603	8	AQ351433	AQ351433 RBC11-11
15	249	13.3	1107	9	CNS04176	AL295764 Tetracodon
16	241	12.9	1197	9	AY398777	AY398777 Mus muscu
17	239.5	12.8	1197	9	AY398776	AY398776 Pan trogl
18	238.5	12.7	1197	9	AY398775	AY398775 Homo sapi
19	231.5	12.4	1086	9	AY399291	AY399291 Homo sapi
20	231.5	12.4	2583	3	CR609171	CR609171 full-length
21	218	11.6	1146	9	AY414479	AY414479 Mus muscu
22	218	11.6	2785	3	AK028281	AK028281 Mus muscu
23	217	11.6	2919	3	AK038620	AK038620 Mus muscu
24	217	11.6	2959	3	AK079529	AK079529 Mus muscu
25	216.5	11.6	2432	3	AK051189	AK051189 Mus muscu
26	216.5	11.6	2708	3	AK076275	AK076275 Mus muscu
27	215.5	11.5	1086	9	AY399293	AY399293 Mus muscu
28	215.5	11.5	2118	3	AK031109	AK031109 Mus muscu
29	215.5	11.5	2292	3	AK036597	AK036597 Mus muscu
30	215.5	11.5	2432	3	AK036690	AK036690 Mus muscu
31	213	11.4	863	5	BP161654	BP161654
32	213	11.4	1511	3	CR601301	CR601301 full-length
33	213	11.4	1603	3	CR614199	CR614199 full-length
34	213	11.4	1612	3	CR594588	CR594588 full-length
35	213	11.4	1619	3	CR610268	CR610268 full-length
36	213	11.4	1621	3	CR614594	CR614594 full-length
37	213	11.4	1626	3	CR598713	CR598713 full-length
38	213	11.4	1629	3	CR596547	CR596547 full-length
39	213	11.4	1629	3	CR623838	CR623838 full-length
40	213	11.4	1640	3	CR619476	CR619476 full-length
41	213	11.4	1642	3	CR594428	CR594428 full-length
42	213	11.4	1650	3	CR605131	CR605131 full-length
43	213	11.4	1682	3	CR614663	CR614663 full-length
44	207.5	11.1	1662	9	AY400424	AY400424 Mus muscu
45	207.5	11.1	1200	9	AY416064	AY416064 Homo sapi

ALIGNMENTS

RESULT 1
CC905917/c
LOCUS
DEFINITION CC905917.1 TAMB1 Bos taurus genomic clone t025011ba, genomic survey sequence.
ACCESSION CC905917
VERSION CC905917.1 GI:33524850
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 803)
AUTHORS Lin, S., Najjar, F.Z., Adelson, D., Gill, C.A. and Roe, B.A.
TITLE Bovine BAC End Sequences from Library TAMB1
JOURNAL Unpublished (2003)
COMMENT Contact: Bruce A. Roe
Advanced Center for Genome Technology
University of Oklahoma Department of Chemistry and Biochemistry
620 Parrington Oval, Room 208, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu

FEATURES

source
1..803
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Angus bull T A M U Shoshone Y6 11519666"
/db_xref="taxon:9913"

/clone="r025011ba"
 /sex="Male"
 /cell_type="Blood"
 /clone_1ib="TAMR"
 /note="Vector: pBeloBAC11, Site 1: HindIII, Site 2:
 HindIII; TAMR Bovine BAC library (Male) produced by Texas
 A&M University, Department of Animal Science."

ORIGIN

Alignment Scores:

Pred. No.:	1,1e-87	Length:	803
Score:	925.00	Matches:	176
Percent Similarity:	92.75%	Conservative:	3
Best Local Similarity:	91.19%	Mismatches:	13
Query Match:	49.41%	Indels:	1
DB:	9	Gaps:	0

US-09-995-225B-16 (1-353) x CC905917 (1-803)

```

QY 162 ProTyrTyrTyrTyrProAsnIleTyrThrGluAspTyrIleSerThr-SerValHisH 181
DB 765 CCTAACAGCTGGGGGCCCAACCCCTGAGATTGAGATTACACAGCAGGTCAATGATCA 706
QY 181 sValIleuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePheI 201
DB 705 TGCTTAATCGAATTCACCTGCTCCCTGTGACTGCGGTACCCCTGTTCCATATTTCTCAT 646
QY 201 eLeuAsnSerIleIleValTyrIleLeuArgArgIleSerAsnPheArgLeuArgIleTyr 221
DB 645 CTGAACATCCATCATTTGTTCACAGCTCAGAGAGAGAAATTTTCGCTCCGTGGCTCA 586
QY 221 rSerThrGlyLeuThrThraIleLeuPheThrIleThrSerIlePheAlaThrIleuTr 241
DB 585 CTCACCGGGGAAACCGCTGCCATCTGTATTACACACCTCATTTTGCCACCTCTG 526
QY 241 PALAProArgIleIleMetIleLeuTyrHisIleuTyrGlyAlaProIleGlnAsnArgTr 261
DB 525 GGCCCAACGATATCATGATCTCTTACCAACCTTTACCGGAGCCCATCCAGAACCGCTG 466
QY 261 pLeuValHisIleMetSerAspIleAlaAsnMetIleuAlaLeuAsnThrAlaIleAs 281
DB 465 GCTGGTGACAGTCATGTCGAGTGTGCCAACATGCTGCTTGGAAACGGCCATCA 406
QY 281 nPhePheLeuTyrCysPheIleSerIleArgPheArgThrMetAlaAlaIleThrLeuTyr 301
DB 405 CTCTTCTCTCATTTGCTTATCAGCAAGGCGTTCGTGCAAGCGCGCTACTCTCA 346
QY 301 sAlaPhePheLeuTyrCysGlnIleGlnProValGlnPheTyrThrAsnHisAsnPheSerI 321
DB 345 GGCCTTCTTCAAGTGCAGAACCAACCGCTTCACTTACACCAACATTAATCTTTCAT 286
QY 321 eThSerSerProTyrPheSerProAlaAsnSerHisCysIleIleuMetLeuValTyrG 341
DB 285 AACAGAGTACCCCTGATCTCAACCGGCACTCAACATGATCAAGATGCTGATGCCA 226
QY 341 nTyrAspIleuArgIleTyrProIleValSerPro 353
DB 225 GTATGACAAATGAGAAACCTATTAAGATATCCCA 189

```

RESUT 2
 CNS022E4
 LOCUS
 DEFINITION
 Tetradon nigriviridis genome survey sequence T7 end of clone
 227D19 of library G from Tetradon nigriviridis, genomic survey
 sequence.
 AL178069
 A1178069.1 GI:7816126
 GSS: genome survey sequence.
 KEYWORDS
 SOURCE
 ORGANISM
 Tetradon nigriviridis
 Tetradon nigriviridis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;

REFERENCE

1 Tetraodontidae; Tetraodontidae; Tetraodon.

AUTHORS

Roest Crollius,H., Jailon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Winkler,P., Brotier,P., Quetier,F.,
 Saurin,W., and Weissenbach,J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetradon nigriviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)

TITLE

JOURNAL
 MEDLINE
 PUBMED
 20296633
 10835645

AUTHORS

2

TITLE

Roest Crollius,H., Jailon,O., Dasilva,C., Ozouf-Costaz,C.,
 Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
 Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigriviridis
 Genome Res. 10 (7), 939-949 (2000)

JOURNAL

20359837
 10899143

REFERENCE

3 (bases 1 to 806)

JOURNAL

Submitted (12-APR-2000) Genoscope - Centre National de Sequenage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetradon nigriviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 Location/Qualifiers
 1. 806
 /organism="Tetraodon nigriviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="227D19"
 /clone_1ib="G"
 /note="Genoscope sequence ID : C0AG227CB10LPI-end : T7"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:	1,04e-53	Length:	806
Score:	606.00	Matches:	111
Percent Similarity:	65.99%	Conservative:	52
Best Local Similarity:	44.94%	Mismatches:	74
Query Match:	32.37%	Indels:	11
DB:	9	Gaps:	2

US-09-995-225B-16 (1-353) x CNS022E4 (1-806)

```

QY 57 ArgArgGlnIleSerSerTyrAsnTyrIleuAlaIleuAlaIleuVal 76
DB 5 AAGACCAAGAGAGCTGATCACTACCTGCGGCGGAGACAGGCTCGCATCTCTCC 64
QY 77 LeuPhePheIleValPheValAspPheLeuGluAspPheIleuAsnMetGlnMet 96
DB 65 CAGCTTCATCATATTTGCTGCGCTTCTGTGAGACGGGCTTTCACACGGGAGTCC 124
QY 97 ProGlnValProAspIleIleGluValIleuIleuPheSerSerIleHisThrSerIle 116
DB 125 CCCGCGCTCTCTTAACATCACTGACGCGCCGCGAGTTTGCGGCAACACGCTCATC 164
QY 117 TrpIleThrValProLeuThrIleAspArgTyrIleAlaValCysHisProLeuTyr 136
DB 185 TGCTCACACCGTCCCTCTACCGTGAACCGCTACGCGGCTGTCACCCCTCTTCAC 244
QY 137 HisThrValSerTyrProAlaArgThrArgIleValIleValSerValTyrIleThrCys 156
DB 245 AGGAGATCACTACCGGCGGAGAGATATCCGGTGTCTGCTGCTGCTGCTGCTG 304
QY 157 PheLeuThrSerIleProTyrTyrTyrProAsnIleTyrThrGluAspTyrIleSer 176
DB 305 CTGACCTCGGCGCTCTCTTCTGTGTGTCCGACATGTGAGAGAACGCCACCGCGC 364

```

QY 177 ThrSerValHisHisValleuIleTrrIleHisCysPheThrValTyrLeuValProCys 196
Db 365 AMGGCCCTGAGCCGCTCTCATCTGTGACCCAMGTACCATCTTCTGCGCCCTGC 424
QY 197 SerIlePhePheIleLeuHisSerIleIleValTyrIleuArg--ArgIysSerAsn 215
Db 425 AGCATTTCTCGTCTCTCAACTTTGTATCATCCMAAGCTGAGCTGGCGGAGAGCGG 484
QY 216 PheArgLeuArgGly-----TyrSerThrGlyIysThr 226
Db 485 CAGCGCAG--CGGGCGAGCGCGGCMCAAGTMCACCGCTCCGCGCGCTGGGAGAGAGC 543
QY 227 ThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeuTrrAlaProArgIleIle 246
Db 544 ACCGCGCATGCTGTGGCCATCACCTCGCTCTGTCTGTGGCGACCCAGAGCGGTG 603
QY 247 MetIleLeuTyrHisIleuTyrGlyAlaProIleGlnAsnArgTrrLeuValHisIleMet 266
Db 604 GTGCTCATCTACCACTGTACGTGTAGTGTGTTCMAAGCGAMTGGCGCGCTCACCTGGCC 663
QY 267 SerAspIleAlaAsnMetIleuAlaLeuLeuAsnThrAlaIleAsnPheLeuTyrCys 286
Db 664 TACGACCTGTCCAAACATGTGGCGCATGTCTCAACAGCGCGGTAAACTTCTCTGTACTGC 723
QY 287 PheIleSerIysArgPheArg 293
Db 724 TTCGTACGACAGCGCTTCCGC 744

RESULT 3
CL137737 824 bp DNA linear GSS 05-JUN-2004
LOCUS CL137737
DEFINITION ISB1-110A2.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110A2,
genomic survey sequence.
ACCESSION CL137737
VERSION CL137737.1 GI:40631372
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 824)
REFERENCE
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
CONTACT: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submls@wustl.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 2
High quality sequence stop: 672.
FEATURES
source
1..824
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:3164"
/clone_lib="ISB1-110A2"
/clone_1lb="ISB1"
/note="Vector: pBelOBAC11, ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN

Alignment Scores:
Pctd. No.: 2.3e-43 Length: 824
Score: 509.00 Matches: 97
Percent Similarity: 59.51% Conservative: 50
Best Local Similarity: 39.27% Mismatches: 69
Query Match: 27.19% Indels: 31
DB: 9 Gaps: 3

US-09-995-225B-16 (1-353) x CL137737 (1-824)

QY 27 PheValProValValTyrTyrSerIleuLeuCysValGlyLeuProAlaAsnIleLeu 46
Db 157 TTTCTTCTTACAGTG-----AACATGGCTG 180
QY 47 ThrValIleIleLeuSerIleuValAlaArgArgGlnLysSerTyrAsnTrrLeu 66
Db 181 ACTGCTGTGACCTGTCAAGGCTGGCAGAAACAAAGAGCTCTTACAGTTACCTG 240
QY 67 LeuAlaLeuAlaAlaAspIleLeuValIlePhePheIleValPheValAspPheLeu 86
Db 241 TTGGCTCTTACCATCTCACAACATCTTGACCAAGATTTCATCTTTTGGCTTATC 300
QY 87 LeuGluAspPheIleLeuAsnMetGlnMetProGlnValProAspIysIleIleGluVal 106
Db 301 CTGCAACAGCAATACCTCCACGCAAGGCGCCAGTACCTTATCCATGTGTGATGTC 360
QY 107 LeuGluPheSerSerIleHisThrSerIleTrrIleThrValProLeuThrIleAspArg 126
Db 361 CTGAGATTCTCTTAATCATCATCTTATTTGGTACGTATACTGACCGTGACCGT 420
QY 127 TyrIleAlaValCysHisProLeuTyrHisThrValSerTyrProAlaArgThrArg 146
Db 421 TATGTGGCATTTATGTATATCTATCCCTACATATTCGCTTTCTTACCAAGCAACCGT 480
QY 147 LysValIleValSerValTyrIleThrCysPheLeuThrSerIleProTyrTrrPrr 166
Db 481 AGAGTCATTTGTCTTGTCTTCTTCATGCTTGTGACCGGTACCTTCTTACTGTGTG 540
QY 167 ProMetIleTrrThrGluAspTyrIleSerThrSerValHisHisValleuIleTrrIle 186
Db 541 AGTATGTTTGAAGGATTCGCGTACCCAGAGATTGTACCTTATCTTACAGTGAGCC 600
QY 187 HisCysPheThrValTyrIleuValProCysSerIlePhePheIleuAsnSerIleIle 206
Db 601 CACTGCTTATATATATCTTCAATCTTCCCTGACACATATTTTGTATTCACAACTGTGCAT 660
QY 207 ValTyrLysLeuArgArgLysSer-----AsnPheArgLeuArgGly 220
Db 661 ATCTACAGGCTGAGGAAAGAAAGAGGCTCCAGAAATTGCGAGCTTCTTAAATGGGC 720
QY 221 TyrSerThrGlyIysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 721 AAGACCCACAGCAAAATTTCCTCCGGGCGCATTAACCTGTGTTTTCGCCGATCTCT 780
QY 241 TrrAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 781 TGG-----GGCTCTCTTAAAAAAACCAT 804
QY 261 -TrrLeuValHisIleMet 266
Db 805 TTGCTCATCATCTTAAAT 823

RESULT 4
AZ513631/c 362 bp DNA linear GSS 05-OCT-2000
LOCUS AZ513631
DEFINITION IM0359L07R Mouse 10kb plasmid UGCGM library Mus musculus genomic
clone UGCGM0359L07 R, genomic survey sequence.
ACCESSION AZ513631
VERSION AZ513631.1 GI:10694947
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausen, A. and Wright, D., Weis, R.
Mouse whole genome scaffolding with paired end reads from 10kb

Db 72 TACTCGCTCCGGCAGCTCGCGGGGCGCGGGCGGCTGTCGACATTCACCGATCTT 131

Qy 270 AAlaamMetLeuAlaLeuLeuAnThrAlaAlaamPheLeuTyTCyPheLeuSer 289

Db 132 GCCAACATGTGGCGTGTGCACACGGGGTTCCTTCTCTACCTGTTTCATCAGC 191

Qy 290 LysArgPheArgThrMetAlaAlaAlaThrLeuValAlaPhePheLysCysGlnLysGln 309

Db 192 WAGCGTTTCGGGGGCGATGGCGGCCAACGCTCTCGAGCGCTGCCATCGACAGAAAGCAG 251

Qy 310 ProValGlnPheTyThrAsnAlaamPheSerIleThrsSerProTyrPilesPro 329

Db 252 CTGCACCGCTTCAACCGCAGCACAACCTTTTCATACGAGTGGCCCTGGTCTCAGCA 311

Qy 330 AAlaamSerHisCysAlaLeuMetLeuValTyGlnTyThrAspLysGlnLysProIle 349

Db 312 GCMACTCCACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371

Qy 350 LysValSer 352

Db 372 TGTATTTC 380

RESULT 6
LOCUS CL137848 867 bp DNA linear GSS 05-JUN-2004
DEFINITION ISB1-110C15.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110C15,
genomic survey sequence.
ACCESSION CL137848
VERSION CL137848.1 GI:40631483
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis

REFERENCE
AUTHORS Eukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 867)
Kremutzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
CONTACT: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submls@wustl.edu
Insert Length: 7500 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 11
High quality sequence stop: 673.
Location/Qualifiers
1. 867
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-110C15"
/clone_1lb="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:
2,07e-36	444.00	61.06%	40.71%	23.72%
Length: 867	Matches: 92	Conservative: 46	Mismatches: 62	Indels: 26
Gaps: 4				

US-09-995-225B-16 (1-353) x CL137848 (1-867)

Qy 27 PheValProValValTyTySerLeuLeuLeuGlyLeuProAlaamIleLeu 46

Db 172 TTTCTTCTTACAGTG-----AACATGCTG 195

Qy 47 ThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLysSerSerTyrrantTyrrleu 66

Db 196 ACTGCTGGACCTCTCAAGGCTGACGAGAAACAAAGCTCTTCACTACCTG 255

Qy 67 LeuAlaLeuAlaAlaAlaPheLeuValLeuPhePheIleValPheValaPheLeu 86

Db 256 TTGGCTCTTACCATCTCAACATCTTGGACCGAGATTTCATCTTTTGGCTTATC 315

Qy 87 LeuGlnaPhePheIleLeuamMetGlnMetProGlnValProaPheIleIleGlnVal 106

Db 316 CTGCMAACAGCATACTCCACCGCAAGGCGCCAGTACCTTAATCATGTGTCAGTGC 375

Qy 107 LeuGlnPheSerSerIleHisThrSerIleTyrrPheThrValProLeuThrIleAspArg 126

Db 376 CTGAGTTCTCTCTTAATGATCATCTTAATTTGGTGACGTATGATACCTGACCGCT 435

Qy 127 TyrrIleAlaValCysHisProLeuTyrrHisThrValSerTyrrProAlaArgThrArg 146

Db 436 TATGGCATTTATGCTATTCCTGCTACATATTCGCTCTTCTTACCCAGACGACCTGT 495

Qy 147 LysValIleValSerValTyrrIleThrCysPheLeuThrSerIleProTyrrTyrrP 166

Db 496 AGAGTCATTTGTCTTGTCTTCCCTCATGCTTTGGACCGGTACCTTCTACCTGTG 555

Qy 167 ProLeuIleTyrrThrGlnaPheTyrrIleSerThrSerValHisValLeuIleTyrr 186

Db 556 AGTGATGTTTGAGGAGGATCGGCTACCCAGATTCGTACCTTAATCAAGTGACC 615

Qy 187 HisCysPheThrValTyrrleu--ValProCysSerIlePhePheIleLeuamSerIle 206

Db 616 CACTGCTTTATATATATCTTCCATTCCTCGACCATATTTTGGATTAACCAACTCG 675

Qy 206 LysValTyrr-----LysLeuArgArgLysSer-----Asp 216

Db 676 GTCATTATCTTACGACGCTGAGGAAACAAACCGGGGTCCCAAGAGTTGCCATCT 735

Qy 216 heArgLeuArgGlyTyrrSerThrGlyLysThrAlaIleLeuPheThrIleThrSer 236

Db 736 TTCTTAAATGGGCAAGAA-----CCGACGATCAATTTCTTCCCTTGAGACAA 789

Qy 236 IePheAlaThrleu 240

Db 790 TTTACCCCTCCTG 803

RESULT 7
LOCUS BZ169352 813 bp DNA linear GSS 11-OCT-2002
DEFINITION CH230-259E22.TU CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-259E22, genomic survey sequence.
ACCESSION BZ169352
VERSION BZ169352.1 GI:23810547
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE
AUTHORS Eukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 813)
Zhao,S., Shetty,J., Shatsman,S., Tsagaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D.,
Riggs,F., de Jong,P. and Frazer,C.M.
Rat BAC End Sequences from Library CHORI-230 MDOI segment
Unpublished (1999)
Other GSSs: CH230-259E22.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library availability, please contact Pieter de Jong (pdejong@gmail.com). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/or> ering information.htm). BAC end plate: http://www.tigr.org/tcd/bac_ends/rat/bac_end_intro.html Seq primer: SP6
 Class: BAC ends.

FEATURES

source location/Qualifiers
 1. 813
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SanHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-259E22"
 /sex="Female"
 /cell_type="Brain"
 /clone_id="CHORI-230 Segment 2"
 /note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI; CHORI-230 Rat (BN/SanHsd/MCW) BAC library produced by Pieter de Jong"

ORIGIN

Alignment Scores:

Pred. No.: 2.48e-35 Length: 813
 Score: 433.50 Matches: 90
 Percent Similarity: 88.57% Conservative: 3
 Best Local Similarity: 85.71% Mismatches: 5
 Query Match: 23.16% Indels: 7
 DB: 8 Gaps: 1

US-09-995-225b-16 (1-353) x BZ169352 (1-813)

```

QY      37  LeuCysleuGly-----LeuProAlaasnileuThrValIle 49
DB      499  CTTCGCTTGGGATAACTTGAATCTCTCTCTCCACGACAAATCTTGCACAGTCATC 558
QY      50  IleLeuSerGlnleuValAlaArgArgGlnIlySerSerTyAsnTyIleuValleu 69
DB      559  ATCTCTCTCAACTGTGTACCCAGAAAGAGTCTCTTACACATCTTGTGCACATT 618
QY      70  AlaAlaAlaAspIleleuValIleuPhePheIleValPheValAspPheleuGluasp 89
DB      619  GCGTGGCCGACATCTTGCTCTCTTTTCATCGCTTGTGATTTCTGTTCAGAACAC 678
QY      90  PheIleleuAsnMetGlnMetProGlnValProAspIlyIleIleGluValleuGluPhe 109
DB      679  TTCAATTTGACCATGACATGCGCTCCGATCCCTGACAGATCATGAAGTCTAGAGTTC 738
QY      110 SerSerIleHisThrSerIleThrIleThrValProleuThrIleAspArgTyIleAla 129
DB      739  TCTTCATTCACACTCTTAATTTGGATTCAGATCCCTTAACAGTTGATGATATATCGCA 798
QY      130 ValGlyHisProleu 134
DB      799  GTCTGTCAACCCACTC 813

```

RESULT 8
 CNS02ENP
 LOCUS
 DEFINITION
 Tetradon nigraviridis genome survey sequence PUC-ori end of clone
 262F08 of library G from Tetradon nigraviridis, genomic survey
 sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 GSS: genome survey sequence.
 Tetradon nigraviridis
 Tetradon nigraviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE

1

AUTHORS
 Roest Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fitzmes, C., Wincker, P., Brotier, P., Queller, F.,
 Sautin, W., and Weissenbach, J.
 TITLE
 Estimate of human gene number provided by genome-wide analysis
 using Tetradon nigraviridis DNA sequence
 JOURNAL
 Nat. Genet. 25 (2), 235-238 (2000)
 MEDLINE
 20296633
 PUBMED
 10835645
 REFERENCE
 2
 Roest Crolius, H., Jallion, O., Dasilva, C., Ozouf-Costaz, C.,
 Fitzmes, C., Fischer, C., Bouneau, L., Billault, A., Queller, F.,
 Sautin, W., Bernot, A., and Weissenbach, J.
 TITLE
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetradon nigraviridis
 JOURNAL
 Genome Res. 10 (7), 939-949 (2000)
 MEDLINE
 20359837
 PUBMED
 10899143
 REFERENCE
 3
 (bases 1 to 898)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetradon nigraviridis
 genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.
 FEATURES
 source location/Qualifiers
 1. 898
 /organism="Tetradon nigraviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="262F08"
 /clone_1lb="G"
 /note="Genoscope sequence ID : C0AG262D004SP1-end :
 PUC-ori"

ORIGIN

Alignment Scores:
 Pred. No.: 1.79e-34 Length: 898
 Score: 426.00 Matches: 76
 Percent Similarity: 67.65% Conservative: 39
 Best Local Similarity: 44.71% Mismatches: 55
 Query Match: 22.76% Indels: 1
 DB: 9 Gaps: 0

US-09-995-225b-16 (1-353) x CNS02ENP (1-898)

```

QY      42  ProAlaAsnIleuThrValIleIleuSerGlnleuValAlaArgArgIlySer 61
DB      373  CCAGTAAACATCTTGAACGCGGTGGCGTGGACCGCTGCGCCGACCAAGAACGA 432
QY      62  SerTyAsnTyIleuValAlaAlaAlaAspIleleuValIleuPhePheIleVal 81
DB      433  CTGTACTACTACTGCTGCGGTGACAGCGCTCGACATCTCTCCACACTCTTCAATC 492
QY      82  PheValAspPheleuGluaspPheIleleuAsnMetGlnMetProGlnValProasp 101
DB      493  TTGCTGGGCTTCTCTGTGGAGCGGACAGTTCACCGGAGGTCGCCGCTCTCTTA 552
QY      102 LysIleIleGluValleuGluPheSerSerIleHisThrSerIleThrValPro 121
DB      553  CACTCAGTCAGCGCGCGCGGATTTGCCCAACACACGCGCTCATGTGTCCACGTC 612
QY      122 LeuThrIleAspArgTyIleAlaValGlyHisProleuTyTyHisThrValSerTy 141
DB      613  CTCACCGTGACCGGTAGCGGTGCGGTGCGCACCCCTCCACACAGACATCAGCTAC 672
QY      142 ProAlaAspGlyThrValIleValSerValTyIleThrTyPheleuThrSerIle 161
DB      673  CCGGCGCGGACACGAGATATCGCGGTGCTGTGGCTGTGGCTCGGCGCTG 732
QY      162 ProTyTyTyTyTyProAsnIleTyThrGluAspTyIleSerThrSerValHis 181

```

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT	FEATURES	ORIGIN	ALIGNMENT SCORES	SCORE	PERCENT SIMILARITY	BEST LOCAL SIMILARITY	QUERY MATCH	DB	
793	CCCTTCTCTGTGTGTGTCGACATGTGGAGGAACACCCGCGGACGCGS-CTGAGACGC	791	182	ValleuilerPlethieCyapherTherValTyreuleuValProCySerlePhephele	201	792	GTCCTCATCTGAGACCACTGACATCATCTACTCTTCCGCTGAGCATCTTCTGTGTC	851	202	LeuanserilleleValTyryleuArg 211	852	CTCAACTCTTGATCATCCACACGCTGAGG	881	US-09-995-225B-16 (1-353) x AL921815 (1-515)	1.36e-33	Length: 515	414.50	Matches: 104	37.46%	Conservative: 14	Mismatches: 29	Indels: 168	Gaps: 4
177	CTGCCAGCAACATCTCTGACAGTATCTCTTCTTCTGCTG	218	41	LeuProAlaenilleleTherValilleleuSerGinleuValAlaArgArgGln	60	117	CGCGTGTCTCTGCGACAGTTCACGTCGTCACGTCACGTCCTTACTGCTGCTGCGC	176	21	SerAlaCyGlyleuGlyPheValProValValTyrySerleuLeuCyaleuGly	40	117	CGCGTGTCTCTGCGACAGTTCACGTCGTCACGTCACGTCCTTACTGCTGCTGCGC	176	177	CTGCCAGCAACATCTCTGACAGTATCTCTTCTTCTGCTG	218	177	CTGCCAGCAACATCTCTGACAGTATCTCTTCTTCTGCTG	218	177	CTGCCAGCAACATCTCTGACAGTATCTCTTCTTCTGCTG	218

OY		61	SerSerTyrTrentyTrLeuLeuAlaLeuAlaAlaAlaAspLileuValIleuPhePheIle	80
Dd		218	-----	218
OY		81	ValPheValAapPheLeuLeuGluaApPheIleLeuAmMetGlnMetProGlnValPro	100
Dd		218	-----	218
OY		101	AspLysIleIleGluValIleGluIuPheSerSerIleHisThrSerIleTrpIleThrVal	120
Dd		218	-----	218
OY		121	ProLeuThrIleAspArgTyrlIeAlaValCyeHisProLeuLysTyrlHisThrValSer	140
Dd		218	-----	218
OY		141	TyrProAlaArgThrArgLysValIleValSerValTyrlIleThrCySPheLeuThrSer	160
Dd		218	-----	218
OY		161	IleProTyrlTyrlTrpTrpProAmIleTrpThrgLuAspTyrlIleSerThrSerValHis	180
Dd		218	-----	218
OY		181	HisValIleuIleTrpIleHisCySPheThrValTyrlLeuValProCySPserIlePhePhe	200
Dd		218	-----	218
OY		201	IleLeuAmSerIleIleValTyrlLysLeuAlaArgLysSerXanPhe-ArghLeuAGl	220
Dd		219	-----CTCCGGCTTCGAGG	233
OY		220	YTYSerThrGlySthrThrThralalIleLeuPheThrIleThrSerIlePheAlaThrLe	240
Dd		234	ATACTCGACCGGAAGAACCACGGCTATCCTTTTAGCATCATCACCTCACTGTCCGCCATT	293
OY		240	UTPrLaAProArghIleIleMetIleLeuTYRHisLeuTYR-----GlyAlaPr	256
Dd		294	ATGGGCGCGCTCGTACACTATGATTTCTTTATCATCTTTACACGGTGCAACACGAGATGCC	353
OY		256	CILGlnAmArghTrpleuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLe	276
Dd		354	GGGTCCGCGCTAGA---CTCCTGCATTGGTAACGGATGTGCGAACATGCTTGCTACT	410
OY		276	uasnrThralalleasnPheLeuTYRCysPheIleSerLysArghPheArghThrMetAl	296
Dd		411	CAACACCGGGGTCAACTTCTTCTCTACTGCTTTATAGCAAGCGCTTTTCGAGAGATGCC	470
OY		296	aalaIaIaThrLeuLysAlaPhePheLysCySGlnLysGlnPro	310
Dd		471	CGGACGCGTGCTCAAAAGCCTTTTTCCGCTCGCAGAAAACACACCG	513
RESULT 10				
AQ225693/c				
LOCUS				
DEFINITION				
HS_2009_B2_B08_T7_CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2009 Col=16 Row=D, genomic survey sequence.				
ACCESSION				
VERSION AQ225693				
KEYWORDS AQ225693.1 GI:3650922				
SOURCE				
ORGANISM				
Homo sapiens (human)				
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE				
AUTHORS				
Mahatras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T', Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.				
TITLE				
Sequence-tagged connectors: A sequence approach to mapping and sequencing the human genome				
JOURNAL				
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)				

MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L

```

FEATURES
source
location/Qualifiers
1. 448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2009 Col=16 Row=D"
/sex="male"
/clone_11b="CIN Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC clones in
E-Coli DH10B"

```

Alignment Scores:	
Pred. No.:	2.06e-33
Score:	412.00
Percent Similarity:	95.29%
Best Local Similarity:	90.59%
Query Match:	22.01%
DB:	8
US-09-995-225B-16 (1-353) x AQ225693 (1-448)	
	length: 448
	Matches: 77
	Conservative: 4
	Mismatches: 4
	Indels: 0
	Gaps: 0

ORIGIN

ORIGIN

Db	Qy	Score	Length	Matches	Conservative	Mematches	Indels	Gaps
Db	Qy	199	8,31e-23	314.00	100.00%	100.00%	16.77%	8
Db	Qy	139	1.588	66	0	0	0	0
Db	Qy	83	588	66	0	0	0	0

US-09-995-225B-16 (1-353) x AQ344035 (1-588)

Alignment Scores:

Pred. No.: 8,31e-23

Score: 314.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 16.77%

DB: 8

ORIGIN

Location/Qualifiers

1..588

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="GDB:7547470"

/db_xref="taxon:9606"

/clone="RPC1-11-124023"

/sex="Male"

/cell_type="Lymphocytes"

/clone_1ib="RPC1-11"

/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPB11 Human Male BAC library"

Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tcd/hungen/bac_end_search/bac_end_search.html
 Seq primer: 17
 Class: BAC ends.

Db	79	GTGAGCTTCCTGGTGGAGATTTATCTTGAACATGACGATGCCCTCAGTCCCGACAA	20
Oy	103	lelele uVal leu lu 108	
Db	19	ATCATGAAAGTGCTGGAA 2	
RESULT 14			
LOCUS	AO351433.c		
DEFINITION	AO351433	603 bp	DNA
ACCESSION	AO351433		linear
VERSION	AO351433.1		GSS 07-MAY-1999
KEYWORDS	SSS. GI:4178768		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 603)		
JOURNAL	Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and		
COMMENT	Venter, J.C.		
	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready		
	Map Building		
	Unpublished (1997)		
	Other GSSs: RPCI11-113113.TV		
	Contact: Shaying Zhao, William Niernan, Mark Adams		
	Department of Eukaryotic Genomics		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850		
	Tel: 301 838 0200		
	Fax: 301 838 0208		
	Email: hbe@tigr.org		
	Clones are derived from the human BAC library RPCI-11. For BAC		
	library availability, please contact Pieter de Jong		
	(pieter@jlong.med.buffalo.edu). Clones may be purchased from		
	BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from		
	Research Genetics (info@reagen.com). BAC end search page:		
	http://www.tigr.org/tcdb/human/bac_end_search/bac_end_search.html		
	Seq primer: SP6		
	Class: BAC ends.		
FEATURES	Location/Qualifiers		
source	1..603		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="GDB:7543212"		
	/dx_xref="taxon:9606"		
	/clone="RPCI-11-113113"		
	/sex="Male"		
	/cell_type="Lymphocytes"		
	/clone_11b="RPCI-11"		
	/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;		
	RPCI11 Human Male BAC Library"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.94e-22	Length:	603
Score:	309.00	Matches:	65
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.51%	Indels:	0
Db:	8	Gaps:	0
US-09-995-225B-16 (1-353) x AO351433 (1-603)			
Oy	43	AlAaenlleuThrrValIlelleleuSerGlnleuValAlAaArgGlnlYseSer	62
Db	195	GCAAAATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGAGACAGAGTCTCC	136
Oy	63	TyAaenYrleuLeuAlaLeuAlaAlaAspilleuValleuPhePheIleValPhe	82
Db	135	TACAACATATCTCTTGACACTCGCTGCTGCCACATCTTGTCCTTTTCATAGTGT	76
Oy	83	ValAsPheleuLeuGluAspPheIleuAsnWetGlnMetProGlnValProAspIys	102

Db		75	GTGACATTCCTGGTGAAGATTATTTACTTGAAATGCAGATGCTCCTCAGTCCCGACAAG	16
Oy		103	lellelguValleu	107
Db		15	ATCATAGAAGTCTG	1
RESULT 15				
CNS04L7F				
LOCUS				
DEFINITION				
Accession				
Version				
Keywords				
Source				
Organism				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 18:37:44 ; Search time 57.6327 Seconds

(without alignment)
3524.170 Million cell updates/sec

Title: US-09-995-225b-16

Perfect score: 1872
Sequence: 1 MHTAHLAANSLSMWSPG.....CIKMLVYQYDKXGKPIKXSP 353

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1872	100.0	353	2 Q6DMJ6	Q6DJ66 homo sapien
2	1638	87.5	388	2 Q8TDU8	Q8TDU8 homo sapien
3	1317	70.4	265	2 Q9BYT4	Q9BYT4 homo sapien
4	1277.5	68.2	383	2 Q712L1	Q712L1 fugu rubrip
5	994	53.1	186	2 Q86SP2	Q86SP2 homo sapien
6	771	41.2	371	2 Q7T2L2	Q7T2L2 fugu rubrip
7	737	39.4	315	2 Q86SL3	Q86SL3 homo sapien
8	737	39.4	462	2 Q72601	Q72601 homo sapien
9	737	39.4	1464	2 Q8NGB0	Q8NGB0 homo sapien
10	713	38.1	365	2 Q7TGN9	Q7TGN9 mus musculi
11	262	14.0	54	2 Q80UC8	Q80UC8 mus musculi
12	260.5	13.9	394	2 Q17478	Q17478 caenorhabd
13	250.5	13.4	398	1 TRR_BOVIN	TRR_BOVIN
14	249.5	13.3	398	1 TRR_SHEEP	TRR_SHEEP
15	249.5	13.3	398	1 Q7PYB7	Q7PYB7 anopheles g
16	248	13.2	549	2 Q9VZM5	Q9VZM5 drosophila
17	242	12.9	412	1 TRR_RAT	TRR_RAT
18	241	12.9	393	1 TRR_MOUSE	TRR_MOUSE
19	239	12.8	462	2 Q8IRF3	Q8IRF3 drosophila
20	238.5	12.7	398	1 TRR_HUMAN	TRR_HUMAN
21	238.5	12.7	398	1 AAR4356	AAR4356 homo sapi
22	236	12.6	488	2 Q71R7	Q71R7 drosophila
23	236	12.6	488	2 AAQ11747	AAQ11747 drosophila
24	235.5	12.6	395	1 TRR_CHICK	TRR_CHICK
25	235	12.6	404	2 Q8JFZ7	Q8JFZ7 xenopus lae
26	234	12.5	478	2 Q71R75	Q71R75 drosophila
27	234	12.5	478	2 AAQ11746	AAQ11746 drosophila
28	233.5	12.5	354	2 Q6YST0	Q6YST0 sus scrofa
29	233.5	12.5	354	2 BAD12131	BAD12131 sus scrofa
30	233.5	12.5	354	2 BAD08660	BAD08660 sus scrofa
31	233	12.4	397	2 Q9DDRI	Q9DDRI xenopus lae

32	233	12.4	540	2 Q9W025	Q9W025 drosophila
33	232	12.4	435	2 Q8SMR3	Q8SMR3 drosophila
34	231.5	12.4	355	1 CKR1_HUMAN	P32246 homo sapien
35	231.5	12.4	355	2 Q9WYU8	Q9WYU8 callithrix
36	231.5	12.4	355	2 AAH64991	AAH64991 homo sapi
37	231	12.3	397	2 Q8JFZ5	Q8JFZ5 xenopus lae
38	230.5	12.3	344	2 Q7PHG9	Q7PHG9 anopheles g
39	230.5	12.3	427	2 Q7RTK2	Q7RTK2 anopheles g
40	228	12.2	412	2 Q7LC54	Q7LC54 homo sapien
41	228	12.2	415	2 Q9GZQ4	Q9GZQ4 homo sapien
42	228	12.2	415	2 Q96AM5	Q96AM5 homo sapien
43	228	12.2	415	2 AAH67776	AAH67776 homo sapi
44	226	12.1	352	2 Q88820	Q88820 rattus norv
45	226	12.1	352	2 Q9R297	Q9R297 rattus norv

ALIGNMENTS

RESULT 1	Q6DMJ6	PRELIMINARY;	PRT;	353 AA.
AC	Q6DMJ6;			
DT	01-OCT-2004 (TREMBlrel. 28, Created)			
DT	01-OCT-2004 (TREMBlrel. 28, Last sequence update)			
DT	01-OCT-2004 (TREMBlrel. 28, Last annotation update)			
DE	G protein-coupled receptor 139.			
GN	Name=GPR139;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	Bonner T.I., Nagle J.W., Kauffman D.;			
RL	Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY635179; AAT65818.1; ..			
KW	RECEPTOR.			
SQ	SEQUENCE 353 AA; 40679 MW; CF1088FCES04F85 CRC64;			
Query Match	100.0%; Score 1872; DB 2; Length 353;			
Best Local Similarity	100.0%; Pred. No. 1,1e-121;			
Matches	353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MHTAHLAANSLSMWSPGACGLGFPVYVYSSLLCGLPANILTVIILSQVAROK 60			
DB	1 MHTAHLAANSLSMWSPGACGLGFPVYVYSSLLCGLPANILTVIILSQVAROK 60			
QY	61 SSYNTILAAADIVLPFIVVDPLEDPIINMQPQVDPDKIEVLEFSSIHSTWTV 120			
DB	61 SSYNTILAAADIVLPFIVVDPLEDPIINMQPQVDPDKIEVLEFSSIHSTWTV 120			
QY	121 PLTIIRYAVGPKLKHVTSYPARTKRVSVSYITCFLSIPYMMNPNTWEDYSTVA 180			
DB	121 PLTIIRYAVGPKLKHVTSYPARTKRVSVSYITCFLSIPYMMNPNTWEDYSTVA 180			
QY	181 HYLWICFTVYLVPCISFIILNSIIVYKLRRKSNRLRGVSGTKTALIFTTISFATL 240			
DB	181 HYLWICFTVYLVPCISFIILNSIIVYKLRRKSNRLRGVSGTKTALIFTTISFATL 240			
QY	241 WAPRIIMLYHLYGAPIQRMVLVHMSDIANNALALNTAINFLYCFISKRRPTMAATL 300			
DB	241 WAPRIIMLYHLYGAPIQRMVLVHMSDIANNALALNTAINFLYCFISKRRPTMAATL 300			
QY	301 KAPFKCQKQVQFYNNHNSITSSPWPISPAHSGIKMLVYQYDKXGKPIKXSP 353			
DB	301 KAPFKCQKQVQFYNNHNSITSSPWPISPAHSGIKMLVYQYDKXGKPIKXSP 353			
RESULT 2	Q8TDU8	PRELIMINARY;	PRT;	388 AA.

```

AC Q8TDU8;
RT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Putative G-protein coupled receptor.
GN Name=GPCR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeda S., Kadowaki S., Haga T., Takasu H., Mitaku S.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083594; BAB89307.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 388 AA; 44893 MW; D8DF1CF0CC95D9D8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 1638; DB 2; Length 388;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ANILTVIILSQLVAROKSSVNYLLAALADILVLFIVFVDFLEDFILNMQMPQVDPK 102
DB 78 ANILTVIILSQLVAROKSSVNYLLAALADILVLFIVFVDFLEDFILNMQMPQVDPK 137
QY 103 IIEVLEFSSIHSTIWITVPLTIDRYIAVCHPLKHTVSYPARTRKVIYSVYITCFLTSLIP 162
DB 138 IIEVLEFSSIHSTIWITVPLTIDRYIAVCHPLKHTVSYPARTRKVIYSVYITCFLTSLIP 197
QY 163 YWMWPNITWEDYISTSVHVLWIHCFVTVLVPCSIFFILNSIYYKLRKSNFRLRGYS 222
DB 198 YWMWPNITWEDYISTSVHVLWIHCFVTVLVPCSIFFILNSIYYKLRKSNFRLRGYS 257
QY 223 TGTATALLFTTISIPATLAPRIIMILVHLYGAPLQNRWLVIHMSDIANMLALTAINF 282
DB 258 TGTATALLFTTISIPATLAPRIIMILVHLYGAPLQNRWLVIHMSDIANMLALTAINF 317
QY 263 FLYCISIKRFRMAAATLKAFKCKQKQPVQFTTNHFSITSSPMWISPCISIKMLVYQY 342
DB 318 FLYCISIKRFRMAAATLKAFKCKQKQPVQFTTNHFSITSSPMWISPCISIKMLVYQY 377
QY 343 DKNGKPIKVPSP 353
DB 378 DKNGKPIKVPSP 388

RESULT 3
Q9BYT4 PRELIMINARY; PRT; 265 AA.
ID Q9BYT4;
AC Q9BYT4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Putative transmembrane receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21853304; PubMed=11863363;
RA Octolenghi C., Barbieri M., McElisavey K., Fellous M.;
RT "Novel paralogy relations among human chromosomes support a link
RT between the phylogeny of doublesex-related genes and the evolution of

```

```

RT sex determination."
RL Genomics 79:333-343(2002).
DR EMBL; AJ303165; CAC33085.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor; Transmembrane.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 265 AA; 30999 MW; 8F14368A492395B CRC64;

Query Match
Best Local Similarity 70.4%; Score 1317; DB 2; Length 265;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ANILTVIILSQLVAROKSSVNYLLAALADILVLFIVFVDFLEDFILNMQMPQVDPK 102
DB 14 ANILTVIILSQLVAROKSSVNYLLAALADILVLFIVFVDFLEDFILNMQMPQVDPK 73
QY 103 IIEVLEFSSIHSTIWITVPLTIDRYIAVCHPLKHTVSYPARTRKVIYSVYITCFLTSLIP 162
DB 74 IIEVLEFSSIHSTIWITVPLTIDRYIAVCHPLKHTVSYPARTRKVIYSVYITCFLTSLIP 133
QY 163 YWMWPNITWEDYISTSVHVLWIHCFVTVLVPCSIFFILNSIYYKLRKSNFRLRGYS 222
DB 194 YWMWPNITWEDYISTSVHVLWIHCFVTVLVPCSIFFILNSIYYKLRKSNFRLRGYS 193
QY 223 TGTATALLFTTISIPATLAPRIIMILVHLYGAPLQNRWLVIHMSDIANMLALTAINF 282
DB 194 TGTATALLFTTISIPATLAPRIIMILVHLYGAPLQNRWLVIHMSDIANMLALTAINF 253
QY 283 FLYCISIKRFRPT 294
DB 254 FLYCISIKRFRPT 265

RESULT 4
Q7T2L1 PRELIMINARY; PRT; 383 AA.
ID Q7T2L1;
AC Q7T2L1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE G protein-coupled receptor 142b (Fragment).
OS Name=GPR142b;
GN Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percomorphia; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22985413; PubMed=14623098;
RA Fredriksson R., Hoglund P.J., Giorlam D.E., Lagerstrom M.C.,
RA Schioth H.B.;
RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
RT receptors lacking close relatives."
RL FEBS Lett. 554:381-388(2003).
DR EMBL; AY288414; AAP72123.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.

```

FT NON TER 1 1
 PT NON TER 383 383
 SQ SEQUENCE 383 AA; 42610 MW; 761B0A3B97ECDD1 CRC64;

Query Match 68.2%; Score 1277.5; DB 2; Length 383;
 Best Local Similarity 66.3%; Pred. No. 1.6e-80;
 Matches 250; Conservative 35; Mismatches 67; Indels 25; Gaps 5;

QY 1 MHTHTAHLANSLMSWSPG-----SACGIGFVVVVVYLLCTGLP-ANLITVITIS 52
 DB 6 MEHSHFSALEPSSSWSTGQHPSEATQGCPLPLPIYYSALLCGLPVALNLTIVILIS 65
 QY 53 QLVARRKSSSYNYLLALAAADILVLFPIFVDFLEDFILNMQPVQPDKRIEVEFSS1 112
 DB 66 QLVMRKSSSYNYLLALAAADILVLLIYFVDPILEDPILATPLPSLSNAVOVLEFSS1 125
 QY 113 HTSIWITVPLTIDRYAVCHPLKRYTVSPARTKRVIVSYVITCFITSIPIYVWPMIWE 172
 DB 126 HTSIWITVPLTIDRYAVCHPLKRYTVSPARTKRVIVSYVITCFITSIPIYVWPMIWE 185
 QY 173 -----DYISTSVHNLWICHTVYLVCSIFPIINSITVYL-RRKSNRFL 218
 DB 186 LPQTGGGAGGAGGNRRVAAQQLVMAHCITVYFELCTVFFSLNAVIVHKLRRRRCFRL 245
 QY 219 RGYSTGTTAILEPTTSIFATLWAPRIIMLYHLYGAPIONRM---LVHIMSDIANMLAL 275
 DB 246 RGYSTGTTAILEPTTSIFATLWAPRIIMLYHLYGAPIONRM---LVHIMSDIANMLAL 305
 QY 276 LMTAINFPLFCFSKFRPTMAATLKAFFKOKOPVQPYTNHNSITSSPWISPAHSCHI 335
 DB 306 LMTGVNFFLFCFSKFRPTMAATLKAFFKOKOPVQPYTNHNSITSSPWISPAHSCHI 365
 QY 336 KMLVYQYDKNGKRIKVS 352
 DB 366 KMLVYQYDKNGKRIKVS 382

RESULT 5

Q86SP2 PRELIMINARY; PRT; 186 AA.
 ID Q86SP2
 AC Q86SP2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE G protein-coupled receptor FGR3 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxID=9606;
 RX MEDLINE=22584407; PubMed=12679517;
 RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchoris J.E.,
 Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
 Bergmann J.E., Gallanaris G.A.;
 RT "The G protein-coupled receptor repertoire of human and mouse."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
 DR EMBL: AY255545; AA085057.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS50262; G_PROTEIN_RECPT_P1_2; 1.
 KM Receptor.
 FT NON TER 1 1
 FT NON TER 186 186
 SQ SEQUENCE 186 AA; 21687 MW; E402AB947EC080AF CRC64;

Query Match 53.1%; Score 994; DB 2; Length 186;
 Best Local Similarity 100.0%; Pred. No. 3.3e-61;
 Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 NITWEDYISTSVHNLWICHTVYLVPCSIPIINSITVYKLRKSNRFLRGYSTGKT 227
 DB 1 NITWEDYISTSVHNLWICHTVYLVPCSIPIINSITVYKLRKSNRFLRGYSTGKT 60
 QY 228 AILFTTTSIFATLWAPRIIMLYHLYGAPIONRMVLHNSDIANMLALNTAINFPLVCF 287
 DB 61 AILFTTTSIFATLWAPRIIMLYHLYGAPIONRMVLHNSDIANMLALNTAINFPLVCF 120
 QY 288 ISKRFPTMAATLKAFFKOKOPVQPYTNHNSITSSPWISPAHSCHIQLVYQYDKNGK 347
 DB 121 ISKRFPTMAATLKAFFKOKOPVQPYTNHNSITSSPWISPAHSCHIQLVYQYDKNGK 180
 QY 348 PIKVPSP 353
 DB 181 PIKVPSP 186

RESULT 6

Q7T2L2 PRELIMINARY; PRT; 371 AA.
 ID Q7T2L2
 AC Q7T2L2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE G protein-coupled receptor 142a (Fragment).
 GN Name=GR142a;
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Acanthopterygii; Acanthopterygii; Percormorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Takifugu.
 NC NCB1_TaxID=31033;
 RX MEDLINE=223985413; PubMed=14623098;
 RA Fredriksson R., Hoglund P.J., Gloriam D.E., Lagerstrom M.C.,
 Schioth H.B.;
 RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
 RT receptors lacking close relatives."
 RL FEBS Lett. 554:381-388(2003).
 DR EMBL: AY288413; AAF72122.1; -
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0004872; F:receptor activity; IEA.
 DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_P1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPT_P1_2; 1.
 KM Receptor.
 FT NON TER 1 1
 FT NON TER 371 371
 SQ SEQUENCE 371 AA; 42000 MW; C1A95B619681AB3 CRC64;

Query Match 41.2%; Score 771; DB 2; Length 371;
 Best Local Similarity 47.0%; Pred. No. 1.6e-45;
 Matches 133; Conservative 65; Mismatches 75; Indels 10; Gaps 2;

QY 21 SACGIGFVVVVVYLLCTGLPANIITVITISQLVARRKSSSYNYLLALAAADILVLFPI 80
 DB 32 SACGIGFVVVVVYLLCTGLPANIITVITISQLVARRKSSSYNYLLALAAADILVLFPI 91
 QY 81 VVDFLEDFILNMQPVQPDKRIEVEFSS1HTSIWITVPLTIDRYAVCHPLKRYTVS 140
 DB 92 IFVGFLETAIVPHREVPALLILRSVSAEPAFNAHSIWSAVPLTVDRYVALCHPLLRQIS 151
 QY 141 YPARTRKRVIVSYVITCFITSIPIYVWPMIWEISTSVHNLWICHTVYLVPCSIPI 200
 DB 152 YPARTRKRVIVSYVITCFITSIPIYVWPMIWEISTSVHNLWICHTVYLVPCSIPI 211
 QY 201 IINSITVYKLRKSNRFL-----RGYST-----GKTALFTTTSIFATLWAPRIIMLY 250

Db 212 VLNLSIIHTLRARQROCCSGDGRGPPRRLGKTMALTAITSVSVLAAPRTVVYI 271
 QY 251 HLYGAPLQNRWIVHIMSDIANMLANTINFLYCFISKRRP 293
 Db 272 HLYVSSVHDMRWVHLAYDLNMLAMLTAVNFFLYCFVSKPPR 314

RESULT 7

Q86SL3

ID 086SL3 PRELIMINARY; PRT; 315 AA.

AC 086SL3;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE G protein-coupled receptor PGR2 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22584407; PubMed=12679517;
 RA Vasiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
 RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
 RA Bergmann J.E., Galanakis G.A.,
 RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
 DR EMBL; AY255622; A085134.1; -.
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 KW Receptor.
 FT NON TER 1 1
 FT NON TER 315 315
 SQ SEQUENCE 315 AA; 35173 MW; C82593E331EB9023 CRC64;

Query Match 39.4%; Score 737; DB 2; Length 315;
 Best Local Similarity 46.3%; Pred. No. 3.1e-43;
 Matches 133; Conservative 57; Mismatches 95; Indels 2; Gaps 2;

QY 21 SAGGLGFVPPVYVYLLCLGLPANILTVIISQVLRQKSSYNVLAALAAADILVLF 80
 Db 30 SPCVAGVIVYVYVYVLLGLFVSLTVAALRLATRRPSYYLLATASDIIQVVI 89
 QY 81 VVDFLLDEFLIMQMPQVDPKIEVLEFSSIHISITVPLTDRIYAVCHPLKHTVS 140
 Db 90 VFAGFLQGAVALARQVPAVVRTANILFEPAANHASVWIAILLTVDRYALCHPLHRAAS 149
 QY 141 VPARTRKIVSVYITCFITSPYWPWNIWTEDYISTSVHVLIMWICFTYVLPSCSIF 200
 Db 150 SPGRTRRAIAVLAALITGIPFYWMLDMMDTSPRLDVLKMAACLIVYFPCGVFL 209
 QY 201 IINSIIVYLRKRSNFRLGYSTGKTTAIIFFTISIPATIMAPRIIMILYHAGAPIONR 260
 Db 210 VTNNSAIHRLRRRSGSGQP-RVGSKTALILGITTLLFLLMAPRVFVWLXMYVAAPVHD 268
 QY 261 WLVIHMSDIANMLALNTAINFLYCFISKRRF-TMAAATLKAFFKC 306
 Db 269 WRVHLALDVANVMVAMLHTAANFGLYCFVSKTFRATVQVHDAVLPFC 315

RESULT 8

Q72601

ID 072601 PRELIMINARY; PRT; 462 AA.

AC 072601;
 DT 01-OCT-2003 (TReMBLrel. 25, Created)
 DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE G protein-coupled receptor 142.
 GN Name-GPR142;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22985413; PubMed=14623098;
 RA Fredriksson R., Hoglund P.J., Gloriam D.E., Lagerstrom M.C.,
 RA Schlotz H.B.;
 RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
 RT receptors lacking close relatives";
 RL FEBS Lett. 554:381-388(2003).
 DR EMBL; AY284421; AAP2130.1; -.
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 KW Receptor.
 FT NON TER 1 1
 FT NON TER 462 462
 SQ SEQUENCE 462 AA; 51106 MW; E9120B52372C75CA CRC64;

Query Match 39.4%; Score 737; DB 2; Length 462;
 Best Local Similarity 46.3%; Pred. No. 4.3e-43;
 Matches 133; Conservative 57; Mismatches 95; Indels 2; Gaps 2;

QY 21 SAGGLGFVPPVYVYLLCLGLPANILTVIISQVLRQKSSYNVLAALAAADILVLF 80
 Db 151 SPCVAGVIVYVYVYVLLGLFVSLTVAALRLATRRPSYYLLATASDIIQVVI 210
 QY 81 VVDFLLDEFLIMQMPQVDPKIEVLEFSSIHISITVPLTDRIYAVCHPLKHTVS 140
 Db 211 VFAGFLQGAVALARQVPAVVRTANILFEPAANHASVWIAILLTVDRYALCHPLHRAAS 270
 QY 141 VPARTRKIVSVYITCFITSPYWPWNIWTEDYISTSVHVLIMWICFTYVLPSCSIF 200
 Db 271 SPGRTRRAIAVLAALITGIPFYWMLDMMDTSPRLDVLKMAACLIVYFPCGVFL 330
 QY 201 IINSIIVYLRKRSNFRLGYSTGKTTAIIFFTISIPATIMAPRIIMILYHAGAPIONR 260
 Db 331 VTNNSAIHRLRRRSGSGQP-RVGSKTALILGITTLLFLLMAPRVFVWLXMYVAAPVHD 389
 QY 261 WLVIHMSDIANMLALNTAINFLYCFISKRRF-TMAAATLKAFFKC 306
 Db 390 WRVHLALDVANVMVAMLHTAANFGLYCFVSKTFRATVQVHDAVLPFC 436

RESULT 9

Q8NGB0

ID 08NGB0 PRELIMINARY; PRT; 1464 AA.

AC 08NGB0;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Seven transmembrane helix receptor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
 RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.,
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB065914; BAC06129.1; -.
 DR Genew; HGNC:20088; GPR142.
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.

```

DR GO: 0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: 0007186; P:G-protein coupled receptor protein signal; IEA.
DR InterPro: IPR00276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_REC_P1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_REC_P1_2; 1.
KW Receptor; Transmembrane.
KW SEQUENCE 1464 AA; 156507 MW; FFGC63FB35ACF7E CRC64;

Query Match 39.4%; Score 737; DB 2; Length 1464;
Best Local Similarity 46.3%; Pred. No. 1,2e-42;
Matches 133; Conservative 57; Mismatches 95; Indels 2; Gaps 2

OY 21 SAGGLFVPPVYYSLLCGLGPANILTVIILSQLVARROKSSYNLLALAAADIVLFFI 80
DB 1113 SPCVAGVPIPIYYSVLIGLSPVSLITFVALARLAFRRRSPSYLLALTASDIITQVI 1172
OY 81 VFVPLLEDFILNMOPQVDPKILIVLSPSSHTSIWTVPLTIRYIAVCHPLKTHVS 140
DB 1173 VFAGLLQGAVALARQVPAVAVRTANILEPAAHSAWVIALITLVDRYALCHPLHRAAS 1232
OY 141 YPARTRKIVSYVITCFPLTSIPYYMNPINTEDYSTSVHVLWIGCFVYVPCSIFF 200
DB 1233 SPGRRRRIAAVLSAALITGIPFYWLDMPKDTSPRLIDEVLKAKCLTYIFICGFLL 1292
OY 201 ILNSIIVYKLRKRSNRLRGVSTGKTTAILETITSIFATLVAAPRIIMILYLGAPIQNR 260
DB 1293 VTNSAIIHRLRRRGSGLOP-RVGSSTAILLGITTLFTLLMAPRVFWMLYHYVAVPHRD 1351
OY 261 WLVHMSDIANWLLALNTRINPFLYCFPSKRR-TMAAATLKAPKFC 306
DB 1352 WRVHLADLVANWVAMVMTATNANFGLYCFKSPKFRATRVGVHDVADLC 1398

RESULT 10
O7TON9 PRELIMINARY; PRT; 365 AA.
ID O7TON9
AC O7TON9;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE G protein-coupled receptor 142.
GN Name=Gpr142;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus;
CX NCBI_Txid=10090;
KN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=2298413; PubMed=14623098;
RA Fredriksson R., Hoglund P.J., Gloriam D.E., Lagerstrom M.C.,
RA Schiöth H.B.;
RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
RL receptors lacking close relatives."
RL FEBS Lett. 554:381-388(2003).
DR EMBL: AY288428; AAP72137.1; -.
DR GO: 0016021; C:integral to membrane; IEA.
DR GO: 0004872; F:receptor activity; IEA.
DR GO: 0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: 0007186; P:G-protein coupled receptor protein signal; IEA.
DR InterPro: IPR00276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G_PROTEIN_REC_P1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_REC_P1_2; 1.
KW Receptor.
KW SEQUENCE 365 AA; 40758 MW; ACA8649245CDAB47 CRC64;

Query Match 38.1%; Score 713; DB 2; Length 365;
Best Local Similarity 46.8%; Pred. No. 1,6e-41;
Matches 138; Conservative 51; Mismatches 96; Indels 10; Gaps 4;

21 SAGGLFVPPVYYSLLCGLGPANILTVIILSQLVARROKSSYNLLALAAADIVLFFI 80

```

Dd		: :: : :	: : : : :::	64	SPECVAGIIPVIYYSVLSTLGP-----VALARLAARTKRPSTRHYLLALNBSDIVTOYII	117
Oy		VFVDLFLEDFILNMQMPDVKILEVFSSHSITSIMITVPETIDRYIAVCPLKYHTVS		81		140
Dd		VFGVEFLQGAVLARQVAVRRTANILLEPANAHSVIAIVLPFTDRYNALCPLRHRATS		118		177
Oy		YPARTRKVYSVTTCCLSTSIPYYWMENITEDEYSTSVHNLWLMHCFFVYLPGCIF		141		200
Dd		SPGTHRAIAAIVGVTLTGTPIPFYMWDVWRADDPSTMDKLTKMACLVLYEPCNVFL	:	178		237
Oy		IINSIIYKKARKSNFLRGSGTGCTALLFTISIFATLMAPRIIMILYHLXGAPIQR	:	201		260
Dd		VTSNAIIDRLKRQQGRGLRPL-VKSSTAILLGGVSLPALMAPRIIMLYMLYHALVARHD	:	238		296
Oy		MVHMVIDIANMIALNTAINFEPLYCPSIKFR-TMAAA TLKAFFKC--OKOPVO	:	261		312
Dd		MRVHALDIANNMLAMLTENVFGIXCFISTFPATRVGOICDHMCACALSQPQO	:	297		351
 RESULT 11 Q80UC8 PRELIMINARY; PRT; 54 AA.						
ID AC Q80UC8		PRELIMINARY;	PRT;	54	AA.	
DT DT 01-JUN-2003 (TREMBLrel. 24, Created)						
DT DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)						
DT DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)						
DE DE G protein-coupled receptor PCR3 (Fragment).						
OS OS Mus musculus (Mouse).						
OC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
OX OX NCBI_TaxId=10090;						
RN RN [1]						
RP RP SEQUENCE FROM N.A.						
RX RX MEDLINE=22584407; PubMed=12679517;						
RA RA Vasiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.B.,						
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller U.R., Wright A.C.,						
RA Bergmann J.E., Galenakis G.A.;						
RT RT "The G protein-coupled receptor repertoires of human and mouse.";						
RL RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003) .						
DR DR EMBL: AY255548; OA080560.1;"-						
DK DK GO:GO:0004872; F:receptor activity; IDA.						
KW KW Receptor.						
FM FM NON TER	1	1				
FT FT NON TER	54	54				
SO SQ SEQUENCE	54	AA; 6082 MW; 5B7BA4A140EF8FD CRC64;				
Query Match	14.0%;	Score 262; DB 2; Length 54;				
Best Local Similarity	94.4%; Pred. No. 4.7e-11;					
Matches	51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;					
Oy	226 TTALTPTTTSIPATLMARPRIIMILYHLXGAPIQRMWLVMNSDIANNMALNTA	279				
Dd	1 TTA LTFITTSIFATTMAPRIIMILYHLXGAPIQRNPWLVM LDVANMTALA NTNA	54				
 RESULT 12 Q17478 PRELIMINARY; PRT; 394 AA.						
ID AC Q17478		PRELIMINARY;	PRT;	394	AA.	
DT DT 01-NOV-1996 (TREMBLrel. 01, Created)						
DT DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)						
DT DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)						
DE DE Hypothetical protein B0334.6.						
CN CN Name=B0334.6;						
CS CS Caenorhabditis elegans.						
OC OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;						
CC CC Phabiditidae; Pelioderinae; Caeonorhabditis.						
OX OX NCBI_TaxId=6239;						
RN RN [1]						
RP RP SEQUENCE FROM N.A.						
RC RC STRAIN=Bristol NZ;						
RX RX MEDLINE=99069613; Pubmed=9851916;						

RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RL Investigating biology".
RL Science 282:2012-2018(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Brstol N2;
RA Swindburne J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z66519; CA91374.3; --
DR PIR: T18704; T18704.
DR WormPep; B0334.6; CE30473.
DR GO; GO:0016021; C: integral to membrane, IEA.
DR GO; GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1, 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR Hypothetical protein.
SQ SEQUENCE 394 AA; 45120 MW; E0EC618D653B8627 CRC64;
Query Match 13.9%; Score 260.5; DB 2; Length 394;
Best Local Similarity 23.6%; Pred. No. 3.4e-10;
Matches 90; Conservative 62; Mismatches 151; Indels 79; Gaps 12;
QY 5 HAHLANSSLSMWSPPSACGLFVPPVYVYSLLCGLPANITVILISQVARRKSSYN 64
DB 31 HEHDEQSISIVWMS-----NVAVLPVIT-----ALIGLACNLMLMAVLISNKTARRIPSWN 80
QY 65 YLLALAAADILVLFVFEVDFLEDFILNMOMPOVDF-----KITEVLEF 109
DB 81 LLIATAVDCSLPLIFAT-----LDVTPLSLPSLAFSTSNHFFSRIVLYIRTLAS 130
QY 110 SSIHNSITVTPVLTIDRYAVCHPLKHTVSYPARTRKIVSVYITCELTSPYYWMP-- 167
DB 131 TTYKSSVLLVAVFNIERYLGVCPPLNSHRWTSRNSKAIATAIVLSFCSIQ--WPLA 187
QY 168 --NITWEDYISTSVHNVLI-----WHCTVYLVGCSTPFILNLSIVYKL 210
DB 168 YTTICFESNSNOQYVYVILMSTNRALQYRTMDVSLAFANVLPIIGLLYNSRIIFTL 247
QY 211 RRKSNFRLGYS-----TGKTAILEFTISIFATLWAPRI--IMILYHL 252
DB 248 RRVVDEDSRKYEETKLSQGLIQHDANNRTRBANMLAVVFMLEFCVGPQAPRILYLD 307
QY 253 YGAPLQNMVLIHIMEDIANMLALTAINFLYCFISGRFTMAATLKAFK--KCKQKP 311
DB 308 YG-QYHPKAILLYVC--LSQQLVFLNASINFCLYCVSVSRKRYRLMKQTLKKEFLHKLEGV 364
QY 312 QFYTNHNSITSSPWISPANSH 333
DB 365 PFOINLKQTKSSSAHTVLSLEBH 386
RESULT 13
TRFR_BOVIN
ID TRFR_BOVIN STANDARD; PRT; 398 AA.
AC 046639;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin
receptor).
GN Name=TRHR;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=Holstein;
RX MEDLINE=98151373; PubMed=9492373;
RT Taketa M., Shimada Y., Ikeda A., Sekikawa K.;
RT "Molecular cloning of bovine thyrotropin-releasing hormone receptor
RT gene.";
RL J. Vet. Med. Sci. 60:123-127(1998).
CC -!- FUNCTION: Receptor for thyrotropin-releasing hormone. This
CC receptor is mediated by G proteins which activate a
CC phosphatidylinositol-calcium second messenger system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; D83964; BA24069.1; --
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR InterPro; IPR009144; ThyrotropinRHR.
DR InterPro; IPR002120; Trelhnm_receptor.
DR Pfam; PF00001; 7tm 1, 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PRINTS; PR01846; TRHR_FAMILY.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1;
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Glycoprotein; Transmembrane.
FT DOMAIN 1 28
FT TRANSMEM 29 51
FT DOMAIN 52 61
FT TRANSMEM 62 83
FT DOMAIN 84 99
FT TRANSMEM 100 121
FT DOMAIN 122 144
FT TRANSMEM 145 168
FT DOMAIN 169 193
FT TRANSMEM 194 215
FT DOMAIN 216 266
FT TRANSMEM 267 288
FT DOMAIN 289 296
FT TRANSMEM 297 319
FT DOMAIN 320 398
FT DISULFID 98 179
FT CARBOHYD 3 3
FT 10 10
SQ SEQUENCE 398 AA; 45165 MW; D9AF4B211A5701B8 CRC64;
Query Match 13.4%; Score 250.5; DB 1; Length 398;
Best Local Similarity 23.2%; Pred. No. 1.7e-09;
Matches 83; Conservative 69; Mismatches 118; Indels 87; Gaps 13;
QY 28 VPVVYVYSLLC-LGLPANITVILISQVARRKSSYN-YLLALAAADILVLFVDF 85
DB 26 VVITLVLITICGIGYGNIMVAVNVR--TKMKRTPLNCVLSLAVADLWV----- 75
QY 86 LLEDFILNMOMPOVPDKI-----IEVLEFSSIHNSITVTPVLTIDRYAVCH 132
DB 76 -----VAAGLPNITDSIYGSWYGVGCLITYLGLINASSCSITAFIERIYALICH 129
QY 133 ELKHYVSYARTRKRYIVSYITCELTSPYYW-----PNITW-EDYISTSVHNVLI 186
DB 130 PTAQFLCTFSRRAKIIIFVWA---FTSIYCMMLFPLDLNISTYDAIVVSCGYKISRN 186
QY 187 HCFYVLPVCSIFPLINSIIVYKL-----RRKSN----- 215
DB 187 YSPFIYLMRPGVYVPMILATVLYGFARILFLNPIDPDKNSMMKQDSTHQNKNIN 246
QY 216 -----PRIRGYSIGKTTAILEFTISIFATLWAP--RIIMILYHLYGADPQNMVLI 267
DB 247 SKTSNRYFNSVTSRQVTKMLAVVILPALIMPFRTLVVNVNSFLSSFGQENWFL-- 303

QY 268 DIANMLALNTAINFLYPCISKRTMAATLKAFKQKOPVQPTYNHSITS 324
 DB 304 -FCRICIYLSAINIPVINYLMOSKFR---AARLKCCKOKKQKVEKPANYVALNYS 355

RESULT 14
 TRFR_SHEEP STANDARD; PRT; 398 AA.
 ID TRFR_SHEEP
 AC Q28596;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin receptor).
 GN Name=TRHR;
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Caprinae; Ovis.
 NC NCB1_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=9720075; PubMed=9048604;
 RA Bockmann J., Beckers T.M., Winter C., Wittkowi H., Winterhoff H., Deuel T., Kreutz M.R.;
 RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells is 3,5',-triiodothyronine, thyrotropin-releasing hormone, and pit-1 independent.";
 RT Independent.";
 RL Endocrinology 138:1019-1028(1997).
 CC - FUNCTION: Receptor for thyrotropin-releasing hormone. This receptor is mediated by G proteins which activate a phosphatidylinositol-calcium second messenger system.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X95285; CAA64606.1; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR009144; ThyrotropinRHR.
 DR InterPro; IPR002120; Trhlm_receptor.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODPSN.
 DR PRINTS; PR01846; TRHRFAMILY.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane. Extracellular (Potential).
 FT DOMAIN 1 28
 FT TRANSMEM 29 51
 FT TRANSMEM 52 61
 FT TRANSMEM 62 83
 FT DOMAIN 84 99
 FT TRANSMEM 100 121
 FT TRANSMEM 122 144
 FT TRANSMEM 145 168
 FT TRANSMEM 169 193
 FT TRANSMEM 194 215
 FT TRANSMEM 216 266
 FT TRANSMEM 267 288
 FT TRANSMEM 289 296
 FT TRANSMEM 297 319
 FT TRANSMEM 320 398
 FT DOMAIN 3 3
 FT DISULFID 98 179
 FT CARBOHYD 3 3
 FT CARBOHYD 10 10

SQ SEQUENCE 398 AA; 45088 MW; 375A311D3DD2A61A CRC64;
 Query Match 13.3%; Score 249.5; DB 1; Length 398;
 Best Local Similarity 23.2%; Pred. No. 1.9e-09;
 Matches 83; Conservative 68; Mismatches 119; Indels 87; Gaps 13;

QY 26 VPVVVYSLIC-IG-PANITVITLISQVVAROKSSYN-YLALAAADILVFYFVDF 85
 DB 26 VVTILLVLTICGLGIVGINIVLVVR--TKMRPTKICVLSLAVADMLV----- 75
 QY 86 LLEDPLNMQMPQVDPKI-----IEVLFSSIHRSIWTVPLTIDRYAVCH 132
 DB 76 -----VAAGLPNIDISIGVMWYGCLCTIYQLGIDINSSCSITAFTERIYAICH 129
 QY 133 PLKHTVSPARTKRVYSVYITCFLTSIPYTW-----PNIWT-EDYISTSVHVLIM 186
 DB 130 PIKAQFLCTFSRAKIIIFVWA--FTSYCMWFLLDLNISTYKDAIVSGYKISRN 186
 QY 187 HGFVYLVPCSFPIFLNISTYVKL-----RKSN----- 215
 DB 187 YSPYILMDGVFVYVPMILATVLGFIARKIIFLSPIDSPKENSNTWKNDSTHQNKN 246
 QY 216 -----FRLRGYSTGKTTAIFTTISIPATLMAP-RIIMILYHYGAPIQRMLVHNS 267
 DB 247 SKTSNRYPNSTVSSRKQVTKMLAVVILFALLMPPRTLVVNSFLSPFOENWFL--- 303
 QY 268 DIANMLALNTAINFLYPCISKRTMAATLKAFKQKOPVQPTYNHSITS 324
 DB 304 -FCRICIYLSAINIPVINYLMOSKFR---AARLKCCKOKKQKVEKPANYVALNYS 355

RESULT 15
 QY 07PYB7 PRELIMINARY; PRT; 440 AA.
 AC 07PYB7;
 DT 01-MAR-2004 (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE AGCP12601.
 GN Name=agCG53608; ORFName=ENSANG0000018385;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
 NC NCB1_TaxID=180454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PEST;
 RA Anopheles Genome Sequencing Consortium;
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
 CC - CAUTION: The sequence shown here is derived from an EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
 CC EMBL; AAB01008987; EAA01756.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODPSN.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_1.
 SQ SEQUENCE 440 AA; 50079 MW; 953A62F3B6A9C95 CRC64;

Query Match 13.3%; Score 249.5; DB 2; Length 440;
 Best Local Similarity 23.3%; Pred. No. 2.1e-09;
 Matches 69; Conservative 67; Mismatches 115; Indels 45; Gaps 8;

QY 31 VYVSLICGLPANTIVITLISQVVAROKSSYNVL-ALAAADILVFYFVDFLED 89
 DB 35 VVMTVALIGLIGNIPSWILSR---PQRRSINVLGLACDVTLLTSTVIFGLCAI 91
 QY 90 F-----ILNMQMPQVDPKI---IEVLFSSIHRSIWTVPLTIDRYAVCHPLKHTVSP 142
 DB 92 YHTGVLYVYHNIQIPKISLVVYPLAMIAQNTASVTLTIVTERVAVACHPLRABALCTY 151

Wed Dec 15 11:04:28 2004

us-09-995-225b-16.rup

Page 8

QY 143 ARTRKXVSYVITICELTSLPIYWMNI-----WTEDYIYSTVHH 161
Db 152 GRALYYVGLIVPSILNLPREFWEVLLISTHPDTGLITCYKASDKMKTNETTYIKVTH- 210
QY 182 VLIVHCFEYLYVPCGSFFILNSIIYYKLRKSNRLGSGTK---TAILFTINSIF 237
Db 211 ---WLWVIEFYFIPFSLISEFNLMITROYRFRANKRKOLISREKREIGLMTLVCVIVF 267
QY 238 ATLMAPRIITILYHAGPIQRNRLVHIVMSDIANLALTNALNPALFYLCFISIKRR 293
Db 268 LILCNLPMAMINIVEAEVSYI-----IEVWVKSNLTLTINSVAFEFYIVFGESEKF 318

Search completed: December 14, 2004, 20:25:08
Job time : 59.6327 secs

Db 181 ATGCGGAGCTGACTTCTTCAACATGCGAGTGGCAGGCTGTGCTCAGCCGCTGAGCC 240
Qy 241 CCTGTGCACTGTGTCGCCCCCGGAGCTCCCGTGGGCGCTGTGAGTGTGGGCGGCGAA 300
Db 241 CCTGTGCACTGTGTCGCCCCCGGAGCTCCCGTGGGCGCTGTGAGTGTGGGCGGCGAA 300
Qy 301 GTTCCAGTGGCACTGCAATGCCCTTCAATGTGTCTCACTGTGGCCATGTAATTCACAC 360
Db 301 GTTCCAGTGGCACTGCAATGCCCTTCAATGTGTCTCACTGTGGCCATGTAATTCACAC 360
Qy 361 GCCCTGTGAGCTGTGACCACTACATGAGCGGTGACCTGCGCGGACCTACATGGCCAGC 420
Db 361 GCCCTGTGAGCTGTGACCACTACATGAGCGGTGACCTGCGCGGACCTACATGGCCAGC 420
Qy 421 GTGTACACACGCGGCGACGTGTGCGGCTTGTGTGGGAGTGGCGCGCTGTGACCACTTC 480
Db 421 GTGTACACACGCGGCGACGTGTGCGGCTTGTGTGGGAGTGGCGCGCTGTGACCACTTC 480
Qy 481 TCTGTGCTGTCTTTCTACATCTGACAGCCATGTGTCCACCCGCGGCTAGAGTGGCCAG 540
Db 481 TCTGTGCTGTCTTTCTACATCTGACAGCCATGTGTCCACCCGCGGCTAGAGTGGCCAG 540
Qy 541 ATGCAAGACGAGAACTGCGGACGCGCAAGCTGTGTGATCGCTACGCTAGTGGCCAGCA 600
Db 541 ATGCAAGACGAGAACTGCGGACGCGCAAGCTGTGTGATCGCTACGCTAGTGGCCAGCA 600
Qy 601 CTGGGCAACCTTACGCGGCTGTGTACTCTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
Db 601 CTGGGCAACCTTACGCGGCTGTGTACTCTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
Qy 661 GACCGGAGACACGCGGCGGCTGTGAGCGGACACAGGCTGTGTGGCCACCGGTGAC 720
Db 661 GACCGGAGACACGCGGCGGCTGTGAGCGGACACAGGCTGTGTGGCCACCGGTGAC 720
Qy 721 ACGCAATTTGGGCTCTGACAGCGCACTATCTGATCTGTGTGGGACACGCTCATCATC 780
Db 721 ACGCAATTTGGGCTCTGACAGCGCACTATCTGATCTGTGTGGGACACGCTCATCATC 780
Qy 781 TCGCGAGGAGAGCGCGCTGTGACGACACTACCTGGGCTACTGCACTTTGTGAAGATTTC 840
Db 781 TCGCGAGGAGAGCGCGCTGTGACGACACTACCTGGGCTACTGCACTTTGTGAAGATTTC 840
Qy 841 TCCAACTCTGCGCTTCTCAGAGTTTGTGACCACTTCTACGCTACATGAAC 900
Db 841 TCCAACTCTGCGCTTCTCAGAGTTTGTGACCACTTCTACGCTACATGAAC 900
Qy 901 CAGAGCTTCCCGACAGCTCCAAAGCTGTGATGAAAAAGCTGCTGCGGAGACCGGAC 960
Db 901 CAGAGCTTCCCGACAGCTCCAAAGCTGTGATGAAAAAGCTGCTGCGGAGACCGGAC 960
Qy 961 TGTCTCCCGACACATGTGGGGGTGACAGAGTGTGCGGTAG 1002
Db 961 TGTCTCCCGACACATGTGGGGGTGACAGAGTGTGCGGTAG 1002

RESULT 2
AX665926 1002 bp DNA linear PAT 26-MAR-2003
LOCUS Sequence 3 from Patent WO242458.
DEFINITION AX665926
ACCESSION AX665926
VERSION GI:29290799
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Tian,H., Zhao,J., Chen,J.L., Cutler,G., An,S., Dai,K. and
TITLE Gupte,J.S.
JOURNAL G-protein coupled receptors
Patent: WO 0242458-A 3 30-MAY-2002;
Tularik Inc. (US)
FEATURES Location/Qualifiers

source 1..1002
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 1..1002
/note="unnamed protein product; human G-protein coupled
receptor (GPCR) TGR339"
/codon_start=1
/protein_id="CAD82888.1"
/db_xref="GI:29290800"
/translation="MWSGSMFNGTGLVEEELPACQDLDLGLSLGLVGVGLCY
NALIVLANHSKASMTMDVYFVMAVAVGLSLAVHLIAPSSRWALMSVGEVH
VALDIPVSVSLVMTADLLSLHYLERALPRTYMASVYTRHVCFFWGLALTSF
SLIPYICSHYSTPILCAKQWNEADATLVPFGYVPALATLALVLSRVREDT
PLDDTGTLEPSARHLVATYCTGPMTPHYLLIHTVVISRGKPDHAYLGLHF
VKDPSKLIAFSSFTVPLLYRYMNSFPSSKLQRLMKLPCQDRHCSPDHNGQVLA"

Query Match 100.0%; Score 1002; DB 6; Length 1002;
Best Local Similarity 100.0%; Pred. No. 3,6e-163; Indels 0; Gaps 0;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGAGCTGACAGCTGTGTTCAACGCGACAGGCGTGTGTGAGAGACTGCTGCTCCAG 60
Db 1 ATGTGAGCTGACAGCTGTGTTCAACGCGACAGGCGTGTGTGAGAGACTGCTGCTCCAG 60
Qy 61 GACCTGACAGCTGGGCTGTGCTGTTGTGCTGTGCGGCGCTGTGTGGGCGTGCAGTG 120
Db 61 GACCTGACAGCTGGGCTGTGCTGTTGTGCTGTGCGGCGCTGTGTGGGCGTGCAGTG 120
Qy 121 GGCCTGTGCTCAACAGCCCTGCTGTGCTGCGCAACTTACACAGCAAGCCAGCATAGC 180
Db 121 GGCCTGTGCTCAACAGCCCTGCTGTGCTGCGCAACTTACACAGCAAGCCAGCATAGC 180
Qy 181 ATGCGGAGCTGTACTTTGTCAACATGCGAGTGTGCGAGCTGTGTCTACAGCGCTGCGC 240
Db 181 ATGCGGAGCTGTACTTTGTCAACATGCGAGTGTGCGAGCTGTGTCTACAGCGCTGCGC 240
Qy 241 CCTGTGACACGCTGTGCGCCCCCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGCGAA 300
Db 241 CCTGTGACACGCTGTGCGCCCCCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGCGAA 300
Qy 301 GTTCAAGTGGCACTGTGACATCTCCCTTCAATGTGTCTCACTGTGTGGCCATGTAATC 360
Db 301 GTTCAAGTGGCACTGTGACATCTCCCTTCAATGTGTCTCACTGTGTGGCCATGTAATC 360
Qy 361 GCGCTGTGAGCTGTGACCACTACATGTGAGCGGTGACCTGCGCGGACCTTACATGGCCAGC 420
Db 361 GCGCTGTGAGCTGTGACCACTACATGTGAGCGGTGACCTGCGCGGACCTTACATGGCCAGC 420
Qy 421 GTGTACACACGCGGCGACGTGTGCGGCTTGTGTGGGAGTGGCGCGCTGTGACCACTTC 480
Db 421 GTGTACACACGCGGCGACGTGTGCGGCTTGTGTGGGAGTGGCGCGCTGTGACCACTTC 480
Qy 481 TCTGTGCTGTCTTTCTACATCTGACAGCCATGTGTCCACCCGCGGCTAGAGTGGCCAG 540
Db 481 TCTGTGCTGTCTTTCTACATCTGACAGCCATGTGTCCACCCGCGGCTAGAGTGGCCAG 540
Qy 541 ATGCAAGACGAGAACTGCGGACGCGCAAGCTGTGTGATCGCTACGCTAGTGGCCAGCA 600
Db 541 ATGCAAGACGAGAACTGCGGACGCGCAAGCTGTGTGATCGCTACGCTAGTGGCCAGCA 600
Qy 601 CTGGGCAACCTTACGCGGCTGTGTACTCTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
Db 601 CTGGGCAACCTTACGCGGCTGTGTACTCTCTCCGCTCCGACGAGGAGACACGCGCTTG 660
Qy 661 GACCGGAGACACGCGGCGGCTGTGAGCGGACACAGGCTGTGTGGCCACCGGTGAC 720
Db 661 GACCGGAGACACGCGGCGGCTGTGAGCGGACACAGGCTGTGTGGCCACCGGTGAC 720
Qy 721 ACGCAATTTGGGCTCTGACAGCGCACTATCTGATCTGTGTGGGACACGCTCATCATC 780
Db 721 ACGCAATTTGGGCTCTGACAGCGCACTATCTGATCTGTGTGGGACACGCTCATCATC 780

QY 781 TCGCGAGGAGAGCCCGTGAAGCAGCACTAAGTGGGGCTAAGCACTTTGGAAGATTTTC 840
DB 781 TCGCGAGGAGAGCCCGTGAAGCAGCACTAAGTGGGGCTAAGCACTTTGGAAGATTTTC 840
QY 841 TCCAACTCCTGGCTTCTTCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
DB 841 TCCAACTCCTGGCTTCTTCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
QY 901 CAGAGCTTCCCGAGCAGAGCTCCAGCGCTGATGAAAAAGCTGCGGGGAGCCGAC 960
DB 901 CAGAGCTTCCCGAGCAGAGCTCCAGCGCTGATGAAAAAGCTGCGGGGAGCCGAC 960
QY 961 TCGTCCCGAGCAGCAGCTGGGGGTGAGCAGAGTCTGCGGTAG 1002
DB 961 TCGTCCCGAGCAGCAGCTGGGGGTGAGCAGAGTCTGCGGTAG 1002

RESULT 3
AX646823 1402 bp DNA linear PAT 04-MAR-2003
LOCUS AX646823 Sequence 1015 from Patent EP1270724.
DEFINITION AX646823
ACCESSION AX646823
VERSION AX646823.1 GI:28799247
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 Suwa, M., Arai, K., Akiyama, Y. and Aburatani, H.
Guanosine triphosphate-binding protein coupled receptors
Patent: EP 1270724-A 1015 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
(UP)

FEATURES
source Location/Qualifiers
1..1402
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
201..1202
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD69537.1"
/db_xref="GI:28799248"
/translation="MNSCSWFGNGTGLVBEIPACODLQLSLSLGLVGVPGVLCY
NALLVLANHSKASMTMPVYFVMAVAGVLSALAPVHLGPPSRMAMVSGEYH
VALQIPFNVSILVAMYSTALLSLDIYERALPRTWASVYVNRHVSGFWMGALITSP
SELFTYICSHVSTRALAEKAKQNAEADATLVFVYVVALATLVYLSRVRREDT
PLDRDTRLEPSARHLVAIVCTQFGLMTPEHYILIGHAVIISRGKPVDAHYGLHF
VDFSKLAFSSSFVTPLRYNNSFPSPKRLRLMKKLPGDHRHCSPDHMGVOOYLA"

ORIGIN
Query Match 100.0%; Score 1002; DB 6; Length 1402;
Best Local Similarity 100.0%; Pred. No. 3.5e-163; Indels 0; Gaps 0;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGACAGTGTTCACAGCAGCAGCGCTGTGAGAGAGTCTGCTGCACG 60
DB 201 ATGTGAGCTGACAGTGTTCACAGCAGCAGCGCTGTGAGAGAGTCTGCTGCACG 260
QY 61 GACCTGCAAGTGGGGCTGTCACTGTTGCTGCTGCTGGGCTGTGTTGGGCGTGCAGT 120
DB 261 GACCTGCAAGTGGGGCTGTCACTGTTGCTGCTGCTGGGCTGTGTTGGGCGTGCAGT 320
QY 121 GGGCTGTGACAAAGCCCGTGTGCTGTGCGCACTTACAGCAAGGCAAGGCAATGAC 180
DB 321 GGGCTGTGACAAAGCCCGTGTGCTGTGCGCACTTACAGCAAGGCAAGGCAATGAC 380
QY 181 ATGCCGAGAGCTGTACTTTGTCAACATGAGAGTGGAGCGCTGTGCTCAAGCGCTGGCC 240
DB 381 ATGCCGAGAGCTGTACTTTGTCAACATGAGAGTGGAGCGCTGTGCTCAAGCGCTGGCC 440

QY 241 CCGTGCACCTGTGCGCCCGCCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGCGAG 300
DB 441 CCGTGCACCTGTGCGCCCGCCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGCGAG 500
QY 301 GTCCAGTGGAGCTGAGATGCCCTTTCATATGTCTCTCACTGTGGCCATGTACTCACC 360
DB 501 GTCCAGTGGAGCTGAGATGCCCTTTCATATGTCTCTCACTGTGGCCATGTACTCACC 560
QY 361 GCGCTGTGAGGCTTGCACCACTTATGAGCTGTGCTGCGCGGACCTTACATGGCAGC 420
DB 561 GCGCTGTGAGGCTTGCACCACTTATGAGCTGTGCTGCGCGGACCTTACATGGCAGC 620
QY 421 GGTACAACAGCGGAGCTGTGCGGCTTGTGTGGGGTGGGCGCTGTGACCACTTC 480
DB 621 GGTACAACAGCGGAGCTGTGCGGCTTGTGTGGGGTGGGCGCTGTGACCACTTC 680
QY 481 TCGTCTGCTCTCTTCTACATCTGACAGCAGTGTCTCCAGCCGCGCTAGAGTGGCCAG 540
DB 681 TCGTCTGCTCTCTTCTACATCTGACAGCAGTGTCTCCAGCCGCGCTAGAGTGGCCAG 740
QY 541 ATGCAGAGAGCAGAGAGCTGCCAGCGCAGCTGTGTTCATGTGCTTACGTGTGCCAGCA 600
DB 741 ATGCAGAGAGCAGAGAGCTGCCAGCGCAGCTGTGTTCATGTGCTTACGTGTGCCAGCA 800
QY 601 CTGGCAGCCCTCTAGCGCGCTGGTGTACTCTCCCGGCTCGGAGGAGAGCAGCGCCCTG 660
DB 801 CTGGCAGCCCTCTAGCGCGCTGGTGTACTCTCCCGGCTCGGAGGAGAGCAGCGCCCTG 860
QY 661 GACCGGAGAGCAGCGCGCGCTGTGAGCGCTGTGAGCAGCAGCGCTGTGAGCAGCGGTGC 720
DB 861 GACCGGAGAGCAGCGCGCGCTGTGAGCGCTGTGAGCAGCAGCGCTGTGAGCAGCGGTGC 920
QY 721 AGCGAGTTGGGCTGTGAGCGCAGCTATCTGATCTGTGTTGGGAGCAGCGTATCATC 780
DB 921 AGCGAGTTGGGCTGTGAGCGCAGCTATCTGATCTGTGTTGGGAGCAGCGTATCATC 980
QY 781 TCGCGAGGAGAGCCCGTGAAGCAGCACTAAGTGGGGCTAAGCACTTTGGAAGATTTTC 840
DB 981 TCGCGAGGAGAGCCCGTGAAGCAGCACTAAGTGGGGCTAAGCACTTTGGAAGATTTTC 1040
QY 841 TCCAACTCCTGGCTTCTTCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
DB 1041 TCCAACTCCTGGCTTCTTCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 1100
QY 901 CAGAGCTTCCCGAGCAGAGCTCCAGCGCTGATGAAAAAGCTGCGGGGAGCCGAC 960
DB 1101 CAGAGCTTCCCGAGCAGAGCTCCAGCGCTGATGAAAAAGCTGCGGGGAGCCGAC 1160
QY 961 TCGTCCCGAGCAGCAGCTGGGGGTGAGCAGAGTCTGCGGTAG 1002
DB 1161 TCGTCCCGAGCAGCAGCTGGGGGTGAGCAGAGTCTGCGGTAG 1202

RESULT 4
BC014241 1776 bp mRNA linear PRI 29-JUN-2004
LOCUS BC014241 Homo sapiens G protein-coupled receptor 146, mRNA (cDNA clone
DEFINITION MGC:20737 IMAGE:456363), complete cds.
ACCESSION BC014241
VERSION BC014241.2 GI:40225566
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheiner, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullehy, S.U., Bosak, S.A., McEwan, P.J.,
McMernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1776)
Strausberg, R.
Direct Submission
Submitted (10-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:15559780.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,
Tusneon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.linl.gov>
Series: IRML plate: 30 Row: J Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19923974.
location/Qualifiers
1..1776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20737 IMAGE:4563636"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_id="NIH_MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..1776
/gene="GPR146"
/note="synonym: PGR8"
/db_xref="locusid:115330"
12..1013
/gene="GPR146"
/codon_start=1
/product="G protein-coupled receptor 146"
/protein_id="AAH14241.1"
/db_xref="GI:15559781"
/db_xref="LOCUSID:115330"
/translation="MSSCSWFGNGTGLVEELPACQDQLGLSLSLGLVGVPGVGLCY
NALVLNLHLSKASMTMPDVFVNMVAAGVGLALAPVHLGLPSSRRMLASVGGVH
VALQIPNVSLVAMYSTALLSLDHYERALPRTYMASVYNTRHVCGFVWGGLLTSF

ORIGIN

Query Match 100.0%; Score 1002; DB 9; Length 1776;
Best Local Similarity 100.0%; Pred. No. 3,4e-163;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTGAGCTGAGAGTGTGTTCAACGGCCAGGGCTGTGTGAGAGAGCTGCTGCTCCAG 60
PDRDNGLEPSAHPLIAYCTCGTMTTPHYLLIGTIVTISRKPDVAHYGLHF
VADPSKLAFSSFTPTILYRKNQSPFSKQRLMKRLLPCDHRHCSPDHMGVOVLA"
12 ATGTGAGCTGAGAGTGTGTTCAACGGCCAGGGCTGTGTGAGAGAGCTGCTGCTCCAG 71
61 GACCTGAGAGCTGAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
72 GACCTGAGAGCTGAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 131
121 GACCTGAGAGCTGAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
132 GACCTGAGAGCTGAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 191
181 ATGCCGAGAGCTGATCTTTGTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
192 ATGCCGAGAGCTGATCTTTGTCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251
241 CCGTGTGACCTGCTGTGAG 300
252 CCGTGTGACCTGCTGTGAG 311
301 GTCCAG 360
312 GTCCAG 371
361 GCCCTGTGAGAGCTGTGAG 420
372 GCCCTGTGAGAGCTGTGAG 431
421 GTGTACAAACAGGGGAG 480
432 GTGTACAAACAGGGGAG 491
481 TCCCTGTGAGAGCTGTGAG 540
492 TCCCTGTGAGAGCTGTGAG 551
541 ATGCAAG 600
552 ATGCAAG 611
601 CTGGCAAG 660
612 CTGGCAAG 671
661 GACCGGAG 720
672 GACCGGAG 731
721 AGCGAG 780
732 AGCGAG 791
781 TCGGAG 840
792 TCGGAG 851
841 TCGAAG 900
852 TCGAAG 911
901 CAG 960
912 CAG 971

QY	961	UGCTCCCGGACCATGGGGGTGAGAGGCTGGGGCTAG	1002
DB	972	TGCTCCCGGACCATGGGGGTGAGAGGCTGGGGCTAG	1013
RESULT 5			
LOCUS	AX330167	1778 bp	DNA
DEFINITION	Sequence 54 from Patent WO0162797.	linear	PAT 11-SEP-2001
ACCESSION	AX330167		
VERSION	AX330167.1	GI:15592185	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Vogel, G., Wood, L.S., Parodi, L.A. and Lind, P.		
TITLE	Novel g protein-coupled receptors		
JOURNAL	Patent: WO 0162797-A 54 30-AUG-2001; PHARMACIA & UPJOHN COMPANY (US)		
FEATURES	Location/Qualifiers		
source	1..1778		
ORIGIN	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		
Query Match	100.0%; Score 1002; DB 6; Length 1778;		
Best Local Similarity	100.0%; Pred. No. 3,4e-163;		
Matches 1002; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGTGGAGCTGAGCTGGTTCAAGGGCAAGGGCTGTTGAGAGAGCTGCTGGCCAGC	60
DB	19	ATGTGGAGCTGAGCTGGTTCAAGGGCAAGGGCTGTTGAGAGAGCTGCTGGCCAGC	78
QY	61	GACCTGAGCTGGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGTGGGCTGTCACTG	120
DB	79	GACCTGAGCTGGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGTGGGCTGTCACTG	138
QY	121	GGCCGTGCTCAACAGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	180
DB	139	GGCCGTGCTCAACAGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	198
QY	181	ATGCCGAGCTGTACTTGTGTCAATGAGCAGTGGCAGGCTGTGTGTGTGTGTGTGTGTGT	240
DB	199	ATGCCGAGCTGTACTTGTGTCAATGAGCAGTGGCAGGCTGTGTGTGTGTGTGTGTGTGT	258
QY	241	CCTGTGCACTGTGCTGGGCCCGGAGCTCCGGTGGGCTGTGTGTGTGTGTGTGTGTGTGT	300
DB	259	CCTGTGCACTGTGCTGGGCCCGGAGCTCCGGTGGGCTGTGTGTGTGTGTGTGTGTGTGT	318
QY	301	GTCACGAGGGCAGCTGCAATCCCTTCAATGTGTCTCACTGTGTGTGTGTGTGTGTGTGT	360
DB	319	GTCACGAGGGCAGCTGCAATCCCTTCAATGTGTGTCTCACTGTGTGTGTGTGTGTGTGT	378
QY	361	GCCCTGTAGAGCTTGACCACTACATGAGCGTGTCACTGCCGAGCTTACATGAGCGAGC	420
DB	379	GCCCTGTAGAGCTTGACCACTACATGAGCGTGTCACTGCCGAGCTTACATGAGCGAGC	438
QY	421	GTTTTCACACAGCGGGCAGTGTGGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	480
DB	439	GTTTTCACACAGCGGGCAGTGTGGGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	498
QY	481	TGCTGTGCTCTTCTTCACTGTGAGCATGTGTCAACCCGCGGCTGTGAGTGGCCAG	540
DB	499	TGCTGTGCTCTTCTTCACTGTGAGCATGTGTCAACCCGCGGCTGTGAGTGGCCAG	558
QY	541	ATGCAAGAGCAAGAGCTGCCGAGCGCAAGCTGTGTTCATGTGTGTGTGTGTGTGTGTGT	600
DB	559	ATGCAAGAGCAAGAGCTGCCGAGCGCAAGCTGTGTTCATGTGTGTGTGTGTGTGTGTGT	618
QY	601	CTGGGACACCTCTTACGAGCTGGTGTACTCTTCCGCGTCCGAGGAGGAGCAAGGCCCTG	660

Db 619 CTGGCACCCCTCTTAAGCGCTGTGTACTCTTCCCGCGTCCGACAGGAGACACGCCCTTG 678

QY 661 GACCGGAGACACGGGCGGCTGAGACCCCTCGGCACACAGGCTGTGTGGCCACCGTGTGC 720

Db 679 GACCGGAGACACGGGCGGCTGAGACCCCTCGGCACACAGGCTGTGTGGCCACCGTGTGC 738

QY 721 ACGAGTTTGGGCTCTGTGACGCCACACTATCTGATCTCTGTGGGGACACCGGTATCATC 780

Db 739 ACGAGTTTGGGCTCTGTGACGCCACACTATCTGATCTCTGTGGGGACACCGGTATCATC 798

QY 781 TCGCGAGGAAAGCCCGTGGACGCAACTACTGTGGGCTACTGCACTTTGTGAGGATTTTC 840

Db 799 TCGCGAGGAAAGCCCGTGGACGCAACTACTGTGGGCTACTGCACTTTGTGAGGATTTTC 858

QY 841 TCCAAACTCTGGGCTTCTCCAGAGCTTTGTGACACCACTTCTCTACCGGTACATGAC 900

Db 859 TCCAAACTCTGGGCTTCTCCAGAGCTTTGTGACACCACTTCTCTACCGGTACATGAC 918

QY 901 CAGAGCTTCCCCAGCAAGCTTCCACCGGTATGAAAAAGCTGCCCTGGGGGACCGGAC 960

Db 919 CAGAGCTTCCCCAGCAAGCTTCCACCGGTATGAAAAAGCTGCCCTGGGGGACCGGAC 978

QY 961 TGCTCCCCCGACACATGGGGGTCCAGAGGTGTCTGGGCTAG 1002

Db 979 TGCTCCCCCGACACATGGGGGTCCAGAGGTGTCTGGGCTAG 1020

AX549394	RESULT 6
LOCUS	AX549394 1785 bp DNA linear PAT 26-NOV-2002
DEFINITION	Sequence 679 from Patent WO02061087.
ACCESSION	AX549394
VERSION	AX549394.1 GI:25814007
KEYWORDS	
SOURCE	
ORGANISM	Homo sapiens (human)
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	
AUTHORS	1
TITLE	Burner,G.C., Roush,C.L. and Brown,J.P. Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides
JOURNAL	Patent: WO 02061087-A 679 08-AUG-2002; Lifespan Biosciences, Inc. (US)
FEATURES	location/Qualifiers
Source	1..1785 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	
Query Match	100.0%; Score 1002; DB 6; Length 1785;
Best Local Similarity	100.0%; Prd. No.3.4e-163;
Matches 1002; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGTGAGACTGACAGCTGGTTCACAGGACAGAGGCTGTGTGAGAGACTGCTGCTGCAG 60
DB	21 ATGTGAGACTGCAGCTGGTTCACAGGACAGAGGCTGTGTGAGAGAGACTGCTGCTGCAG 80
QY	61 GACCTGCAGCTGGGGCTGTCACTGTGTGCTGCTGGGGCTGTGTGTGGGTGTGCCAGTG 120
DB	81 GACCTGCAGCTGGGGCTGTCACTGTGTGCTGCTGGGGCTGTGTGTGGGTGTGCCAGTG 140
QY	121 GAGCTGTGCTACACAGCCCTGCTGATGTGAGCCAACTACACACAGCAGCAGATGACC 180
DB	141 GAGCTGTGCTACACAGCCCTGCTGATGTGAGCCAACTACACAGCAGCAGATGACC 200
QY	181 ATGCTGCAGCTGTACTTTGTCAACATGACAGTGCAGGCTGTGTCTCAAGCCCTTGACC 240
DB	201 ATGCTGCAGCTGTACTTTGTCAACATGACAGTGCAGGCTGTGTCTCAAGCCCTTGACC 260

QY 241 CCTGTGACCTGCTCGGCCCCCGAGCTCCCGGTGGGCGCTGTGGAGTGTGGGGGAGAA 300
 Db 261 CCTGTGACCTGCTCGGCCCCCGAGCTCCCGGTGGGCGCTGTGGAGTGTGGGGGAGAA 320
 QY 301 GTCCACGTGGGCACTGCAATATCCCTTCAATGTGTCTCACTGATGGCCATGTACTCCACC 360
 Db 321 GTCCACGTGGGCACTGCAATATCCCTTCAATGTGTCTCACTGATGGCCATGTACTCCACC 380
 QY 361 GCCCTGTGAGCTCTGACCACTAATGAGGCTGACATGCCGCGGACCTAATATGGCCAGC 420
 Db 381 GCCCTGTGAGCTCTGACCACTAATGAGGCTGACATGCCGCGGACCTAATATGGCCAGC 440
 QY 421 GTGTACAACACGGGCGACAGTGTGGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 480
 Db 441 GTGTACAACACGGGCGACAGTGTGGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 500
 QY 481 TCCCTGCGTCTCTTCAATCTGAGGCAATGTGTCCACCCGCGGCTAAGATGCCCAAG 540
 Db 501 TCCCTGCGTCTCTTCAATCTGAGGCAATGTGTCCACCCGCGGCTAAGATGCCCAAG 560
 QY 541 ATGCAAGACGAGAAAGCTGGCGACGCGACGCTGGTGTTCATCGGCTAACGTTGGTCCAGCA 600
 Db 561 ATGCAAGACGAGAAAGCTGGCGACGCGACGCTGGTGTTCATCGGCTAACGTTGGTCCAGCA 620
 QY 601 CTGGCCACCTCTTACGCGCTGTGTGTACTCTTCCCGGCTCCGACGGAGGACACGCCCTTG 660
 Db 621 CTGGCCACCTCTTACGCGCTGTGTGTACTCTTCCCGGCTCCGACGGAGGACACGCCCTTG 680
 QY 661 GACCGGAGACAGGGGCGGCTGTGAGCCCTGGGACACAGGCTGTGTGGTGGCCACCGTGTGC 720
 Db 681 GACCGGAGACAGGGGCGGCTGTGAGCCCTGGGACACAGGCTGTGTGGTGGCCACCGTGTGC 740
 QY 721 AGCGAGTTTGGGCTCTGAGCGCCACACTATCTGATCTCTGCGGGGACACGGTCAATC 780
 Db 741 AGCGAGTTTGGGCTCTGAGCGCCACACTATCTGATCTCTGCGGGGACACGGTCAATC 800
 QY 781 TCSCGAGGGAAGGCGGCTGTGAGCGCACTACCTGGGCTCTGCACTTTGTGAAGATTTC 840
 Db 801 TCSCGAGGGAAGGCGGCTGTGAGCGCACTACCTGGGCTCTGCACTTTGTGAAGATTTC 860
 QY 841 TCCAAATCTCTGCTCTTCTCCAGCAGCTTGTGACACCACTTCTTACCGCTCACTGAAC 900
 Db 861 TCCAAATCTCTGCTCTTCTCCAGCAGCTTGTGACACCACTTCTTACCGCTCACTGAAC 920
 QY 901 CAGAGTTTCCCGAGGAGCTCCACGCGCTGATGAAAAGCTGCTGCGGGGACCGGAC 960
 Db 921 CAGAGTTTCCCGAGGAGCTCCACGCGCTGATGAAAAGCTGCTGCGGGGACCGGAC 980
 QY 961 TGTCTCCCGGACCATGAGGGGTGACGAGGCTGTGGGCTAG 1002
 Db 981 TGTCTCCCGGACCATGAGGGGTGACGAGGCTGTGGGCTAG 1022

RESULT 7
 AC140819 188370 bp DNA linear HTG 03-MAR-2003
 LOCUS Homo sapiens chromosome 5 clone RP11-734E9, WORKING DRAFT SEQUENCE,
 AC140819 8 unordered pieces.
 ACCESSION AC140819
 VERSION AC140819.1 GI:28631200
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 188370)
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 5
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 188370)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-MAR-2003) Production Sequencing Facility, DOE Joint

COMMENT

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

Project Information
 Center Project Name: 1576465
 Center clone name: RP11-11_734E9

Summary Statistics

Consensus quality: 183379 bases at least Q40
 Consensus quality: 184275 bases at least Q20
 Consensus quality: 184958 bases at least Q20
 Estimated insert size: 175000; agarose-fp estimation
 Estimated insert size: 187670; sum-of-ctnigs estimation
 Quality coverage: 9.36 in Q20 bases; agarose-fp estimation
 Quality coverage: 8.73 in Q20 bases; sum-of-ctnigs estimation.
 NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1155: contig of 1155 bp in length
 * 1156 1255: gap of unknown length
 * 1256 4105: contig of 2850 bp in length
 * 4106 4205: gap of unknown length
 * 4206 10555: contig of 6350 bp in length
 * 10556 10656: gap of unknown length
 * 10656 24895: contig of 14240 bp in length
 * 24896 24995: gap of unknown length
 * 24996 42191: contig of 17196 bp in length
 * 42192 42291: gap of unknown length
 * 42292 60204: contig of 17913 bp in length
 * 60205 60304: gap of unknown length
 * 60305 112109: contig of 51805 bp in length
 * 112110 112209: gap of unknown length
 * 112210 188370: contig of 76161 bp in length.

FEATURES

source
 1..188370
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-734E9"
 /clone_lib="RP11 human BAC library 11"

ORIGIN

Query Match 100.0%; Score 1002; DB 2; Length 188370;
 Best Local Similarity 100.0%; Pred.No.163;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGACGCTGTTCACGACAGGCGTGTGAGAGAGTGCCTGCCAG 60
 Db 46995 ATGTGAGCTGACGCTGTTCACGACAGGCGTGTGAGAGAGTGCCTGCCAG 47054
 QY 61 GACCTGACCTGGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGTGGCGTGCAGTG 120
 Db 47055 GACCTGACCTGGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGTGGCGTGCAGTG 47114
 QY 121 GGCCTGTGTCAACAGCCCTGTGTGTCGCGCACTACAGAGAAAGCCAGATGACC 180
 Db 47115 GGCCTGTGTCAACAGCCCTGTGTGTCGCGCACTACAGAGAAAGCCAGATGACC 47174
 QY 181 ATGCCGAGTGTACTTTGTCAACATGACAGTGTGAGGCTGTGTGAGGCTGTGACC 240
 Db 47175 ATGCCGAGTGTACTTTGTCAACATGACAGTGTGAGGCTGTGTGAGGCTGTGACC 47234
 QY 241 CCTGTGACCTGCTCGGCCCCCGAGCTCCCGGTGGGCGCTGTGGAGTGTGGGGGAGAA 300

Db	47235	CTGTGGACACCTGCTGGGCCCCCGAGAGCTCCCGAGTGGGCGGTGGAGATGTTGGGCGGAGAA	47299
Qy	301	GTCCACGTGGGCACTGCAAGATCCCTTTCAATGTGTCTCACTGGTGGCCATGTACTCCACC	360
Db	47295	GTCCACGTGGGCACTGCAAGATCCCTTTCAATGTGTCTCACTGGTGGCCATGTACTCCACC	47354
Qy	361	GGCCGTGGAGCCCTGGACCACTAATAGAGGTGCACTGGCGGGAGACTTAATAGGGCAGC	420
Db	47355	GGCCGTGGAGCCCTGGACCACTAATAGAGGTGCACTGGCGGGAGACTTAATAGGGCAGC	47411
Qy	421	GTGTACAAACACGCGGCACTGTGTGGGCTTGTGTGGGTGGGTGGCGCTGCTGACCACTTC	480
Db	47415	GTGTACAAACACGCGGCACTGTGTGGGCTTGTGTGGGTGGGTGGCGCTGCTGACCACTTC	47477
Qy	481	TCCTGGCTGCTCTTCTACATCTGACGCAATGTGTCCACCCGCGGCTAGAGTGGCCAG	540
Db	47475	TCCTGGCTGCTCTTCTACATCTGACGCAATGTGTCCACCCGCGGCTAGAGTGGCCAG	47534
Qy	541	ATGAGAAACGAGAAAGCTGCCAGAGCCACGCTGTGTTCATCGGCTACGTTGGTGGACGA	600
Db	47535	ATGAGAAACGAGAAAGCTGCCAGAGCCACGCTGTGTTCATCGGCTACGTTGGTGGACGA	47599
Qy	601	CTGGCCACCCCTCTACGCGCTGTGTCTACTCTCCGCGCTCGGAGGAGAGACAGCCCTG	660
Db	47595	CTGGCCACCCCTCTACGCGCTGTGTCTACTCTCCGCGCTCGGAGGAGAGACAGCCCTG	47655
Qy	661	GACCGGGACACGCGGCGGCTGGAGGACCTCGGACACAGGCTGTGTGTGGCCACGTTGTG	720
Db	47655	GACCGGGACACGCGGCGGCTGGAGGACCTCGGACACAGGCTGTGTGTGGCCACGTTGTG	47711
Qy	721	ACGCAATTTGGGCTCTGGACGCCACACTATCTGATCTCTGTGGGGACACAGTCAATATC	780
Db	47715	ACGCAATTTGGGCTCTGGACGCCACACTATCTGATCTCTGTGGGGACACAGTCAATATC	47774
Qy	781	TCGGAGAGGAAAGCCCGTGGACGCACTACCTGGGGCTATGCACTTTGTAAGATTTC	840
Db	47775	TCGGAGAGGAAAGCCCGTGGACGCACTACCTGGGGCTATGCACTTTGTAAGATTTC	47833
Qy	841	TCCAAATCTCTGGGCTTCTCCAGAGCTTTGTGACACACACTTCTTACCGCTACATGAAC	900
Db	47835	TCCAAATCTCTGGGCTTCTCCAGAGCTTTGTGACACACACTTCTTACCGCTACATGAAC	47899
Qy	901	CAGAGCTTCCCGACGACCTTCCACGCGCTGATGAAAAAGCTGCCCTGCGGGACCGGCAC	960
Db	47895	CAGAGCTTCCCGACGACCTTCCACGCGCTGATGAAAAAGCTGCCCTGCGGGACCGGCAC	47954
Qy	961	TGCTCCCCCGACCATGGGGGTGCAACAGGTGTGGCGTAG	1002
Db	47955	TGCTCCCCCGACCATGGGGGTGCAACAGGTGTGGCGTAG	47996
RESULT 8			
LOCUS	AC073957	196204 bp	DNA linear PRI 09-JAN-2002
DEFINITION	Homo sapiens BAC clone Rpl1-449p15 from 7, complete sequence.		
ACCESSION	AC073957		
VERSION	AC073957.7	GI:17149791	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Sullivan J.B. and Waterston R.		
JOURNAL	Toward a complete human genome sequence		
PUBMED	Genome Res. 8 (11), 1097-1108 (1998)		
REFERENCES	9847074		
AUTHORS	2 (bases 1 to 196204)		
TITLE	Leonard S., Cotton M., Hawkins M. and Spalding L.		
JOURNAL	The sequence of Homo sapiens BAC clone Rpl1-449p15		
REFERENCE	Unpublished (2001)		
	3 (bases 1 to 196204)		

	REFERENCE
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

Waterston, R.H.
Direct Submission
Submitted (07-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 196204)
Waterston, R.H.
Direct Submission
Submitted (29-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 196204)
Waterston, R.H.
Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 196204)
Waterston, R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 29, 2001 this sequence version replaced gi:14589737.

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTR/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Firenza, E., Tatematsu, M., Cataneese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletzer de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.5

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-1244M4; the clone sequenced to the right is RP13-580H13. Actual start of this clone is at base position 1 of RP11-449P15; actual end is at base position 196204 of RP11-449P15.

Data from AC091729 was used to finish this clone, AC073957. Polymorphisms have been identified between AC073957 and AC091729.

Unresolved simple sequence repeats, base positions 26410 to 27787
and 63334 to 64953.

```
FEATURES
    source
        1..196204
            location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="Genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="7"
                /map="7"
            /clone="RP11-449P15"
            /clone_lib="RP11"
            1..278
                /note="match to EST BG992151 (NID:g14396221) "
                203..598
                    /note="match to EST BF814714 (NID:g12147717) "
                    1010..1156
                        /note="match to EST F06563 (NID:g672150) "
                        1122..1414
                            /rpt_family="Alu"
                            1344..1863
                                /note="similar to Homo sapiens EST BF928866 (NID:g12326994) "
                                1932..2224
                                    /rpt_family="Alu"
                                    2189..2393
                                        /note="match to EST F06563 (NID:g672150) "
                                        2270..2636
                                            /rpt_family="L1"
                                            2393..2422
                                                /rpt_family="(CACCC)n"
                                                2745..3367
                                                    /rpt_family="L1"
                                                    4057..4262
                                                        /note="match to EST BI039739 (NID:g14446365) "
                                                        5018..5019
                                                            /note="match to EST AW877271 (NID:g8015626) "
                                                            5068..5455
                                                                /note="match to EST BF087410 (NID:g10893120) "
                                                                5193..5455
                                                                    /note="match to EST AA004851 (NID:g1448598) zh92f02.r1"
                                                                    5405..5761
                                                                        /note="match to EST AW877271 (NID:g8015626) "
                                                                        5829..5833
                                                                            /note="match to EST BG992151 (NID:g14396221) "
                                                                            5844..6140
                                                                                /rpt_family="Alu"
                                                                                6107..6377
                                                                                    /note="match to EST BE073441 (NID:g8420427) "
                                                                                    6136..6426
                                                                                        /note="match to EST BF874571 (NID:g12264701) "
                                                                                        6174..6844
                                                                                            /note="match to EST AU133804 (NID:g10994343) "
                                                                                            7097..7195
                                                                                                /note="similar to Mus musculus EST AA021831 (NID:g1485585) mh82c08.r1"
                                                                                                7108..7195
                                                                                                    /note="match to EST AI200352 (NID:g3752958) qf92h12.x1"
                                                                                                    7108..7195
                                                                                                        /note="match to EST BF516594 (NID:g11601773) "
                                                                                                        7108..7195
                                                                                                            /note="match to EST BG709448 (NID:g13987794) "
                                                                                                            7108..7195
                                                                                                                /note="match to EST T16043 (NID:g518205) "
                                                                                                                7109..7195
                                                                                                                    /note="match to EST AL527751 (NID:g12791244) "
                                                                                                                    7117..7199
                                                                                                                        /note="match to EST AA004851 (NID:g1448598) zh92f02.r1"
                                                                                                                        7117..7195
                                                                                                                            /note="match to EST BF087410 (NID:g10893120) "
                                                                                                                            7141..7195
                                                                                                                                /note="match to EST AI174672 (NID:g6361050) "
                                                                                                                                7150..7195
```

```
misc_feature
    /note="match to EST BG034421 (NID:g12427714) "
    7628..7711
        /note="similar to Mus musculus EST W65144 (NID:g1372795) mh82c08.r1"
        7630..7711
            /note="match to EST AI174672 (NID:g6361050) "
            7630..7711
                /note="match to EST AI200352 (NID:g3752958) qf92h12.x1"
                7630..7711
                    /note="match to EST AL527751 (NID:g12791244) "
                    7630..7711
                        /note="match to EST BF516594 (NID:g11601773) "
                        7630..7711
                            /note="match to EST BG034421 (NID:g12427714) "
                            7630..7711
                                /note="match to EST BG709448 (NID:g13987794) "
                                7630..7711
                                    /note="match to EST T16043 (NID:g518205) "
                                    7630..7711
                                        /note="similar to Mus musculus EST AA021831 (NID:g1485585) mh82c08.r1"
                                        7630..7647
                                            /note="match to EST BF087410 (NID:g10893120) "
                                            7646..7711
                                                /note="match to EST BE855648 (NID:g10368069) "
                                                7687..7711
                                                    /note="match to EST BG768797 (NID:g14079450) "
                                                    7702..7711
                                                        /note="match to EST BI084297 (NID:g14502627) "
                                                        8430..8580
                                                            /note="similar to Gallus gallus EST AU396197 (NID:g7127890) "
                                                            8430..8580
                                                                /note="similar to Mus musculus EST AA087576 (NID:g1630844) mh82c08.r1"
                                                                8430..8580
```

Query Match 100.0%; Score 1002; DB 9; Length 196204;
Best Local Similarity 100.0%; Pred No. 2e-163;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTGAGCTGACAGCTGCTGTTCAACGCGCAGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 60
Db 178576 ATGTGAGCTGACAGCTGCTGTTCAACGCGCAGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 178635
QY 61 GACCTGAGCTGAGGCGTGCATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 178636 GACCTGAGCTGAGGCGTGCATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 178695
QY 121 GGCCTGTGCTACACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 178696 GGCCTGTGCTACACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 178755
QY 181 ATGCGGAGGTGATCTTGTCAACATGAGTGCAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 178756 ATGCGGAGGTGATCTTGTCAACATGAGTGCAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 178815
QY 241 CCTGTGACCTGCTGCGCCCGCAGCTCCCGTGGAGGCGTGTGAGTGTGAGGCGGCGCA 300
Db 178816 CCTGTGACCTGCTGCGCCCGCAGCTCCCGTGGAGGCGTGTGAGTGTGAGGCGGCGCA 178875
QY 301 GTTCACGTGGCCTGCAATGCCCTTCAATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 178876 GTTCACGTGGCCTGCAATGCCCTTCAATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 178935
QY 361 GCCCTGTGAGCTGACACCTGACACCTGACACCTGACACCTGACACCTGACACCTGACACCTGACAC 420
Db 178936 GCCCTGTGAGCTGACACCTGACACCTGACACCTGACACCTGACACCTGACACCTGACACCTGACAC 178995
QY 421 GTGTACAACAGCGGCGAGTGTGCGCTTCTGTGTGAGGCTGCGCGCTGTGACAGCTTC 480
Db 178996 GTGTACAACAGCGGCGAGTGTGCGCTTCTGTGTGAGGCTGCGCGCTGTGACAGCTTC 179055
QY 481 TCCTGCGCTCTTCTCATCTGTGAGGCACTGTGTGCAACCGCGCGCTGAGTGTGCGCAAG 540
Db 179055 TCCTGCGCTCTTCTCATCTGTGAGGCACTGTGTGCAACCGCGCGCTGAGTGTGCGCAAG 179115
```


Db	179056	TCCTGCTGCTCTTCTACATCTGACAGCATGTCATCCCGCGCTAGAGTGCACAG	179115
Qy	541	ATGCAAGAACGAGAAAGCTGCGACGCGACGCTGCTGTTCATGGGCTACGTGGTGCACGA	600
Db	179116	ATGCAGAAACGAGAAAGCTGCCACGACGCTGGTGTTCATCGGCTACGTGGTGCACGA	179175
Qy	601	CTGGCCACCCCTCTACGCGCTGGTGTACTCTCCCGCGTCCGACAGGAGACACGCCCTG	660
Db	179176	CTGGCCACCCCTTACGCGCTGGTGTACTCTCCCGGCTCCGACAGGAGACACGCCCTG	179235
Qy	661	GACCGGGACACGGGCGCGCTGAGCCCTCGGACACAGGCTGCTGTGTGCCACCGTGTGC	720
Db	179236	GACCCGGGACACGGGCGCGCTGAGCCCTCGGACACAGGCTGCTGTGTGCCACCGTGTGC	179295
Qy	721	ACGCAAGTTGGGCTCTGGACGCGCACACTATCTGATCTCTGCTGGGGGACACGGTCATCATC	780
Db	179296	ACGCAAGTTGGGCTCTGGACGCGCACACTATCTGATCTCTGCTGGGGGACACGGTCATCATC	179355
Qy	781	TCGCGAGGGGAAGCCCGCTGACGCGACACTACCTTGGGGCTTACTGCATTTGTGAAGATTTC	840
Db	179356	TCGCGAGGGGAAGCCCGTGAACGCGACACTACCTTGGGGCTTACTGCATTTGTGAAGATTTC	179415
Qy	841	TCCAAACTCTCGGSCCTTCTCCAGCAGCTTTGTGACACCACTTCTCTACACGGCTACATGAAC	900
Db	179416	TCCAAACTCTCGGSCCTTCTCCAGCAGCTTTGTGACACCACTTCTCTACACGGCTACATGAAC	179475
Qy	901	CAGAGCTTCCCAGCAGAGCTCCAAAGGCTGATGAAGAAACCTGCCCTGCGGGGACCGGGCAC	960
Db	179476	CAGAGCTTCCCAGCAGAGCTCCAAAGGCTGATGAAGAAACCTGCCCTGCGGGGACCGGGCAC	179535
Qy	961	TGCTCTCCCGGACCACTGGGGGTGCAGCAGGTGTCTGGCTAG	1002
Db	179536	TGCTCTCCCGGACCACTGGGGGTGCAGCAGGTGTCTGGCTAG	179577

RESULT	9
AR242245	
LOCUS	1614 bp DNA linear PAT 20-DEC-2002
DEFINITION	Sequence 2 from patent US 6472173.
ACCESSION	AR242245
VERSION	AR242245.1 GI:27288068
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1614)
TITLE	Ford J. and Yeung G. Chemokine receptor obtained from a cDNA library of fetal liver-spleen Patent: US 6472173-A 2 29-OCT-2002;
JOURNAL	Location/Organisms
FEATURES	1..1614
Source	/organism="unknown" /mol_type="genomic DNA"
ORIGIN	

Query Match	99.9%	Score 1000.8	DB 6	Length 1614
Best Local Similarity	99.7%	Pred. No. 5.5e-163		
Matches	999	Conservative	3	Mismatches 0; Indels 0; Gaps 0
Qy	1	ATGTGAGACTGCACGCTGGTTCAACGGACAAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG	60	
Db	129	ATGTGAGACTGCACGCTGGTTCAACGGACAAGGGCTGTGTGAGAGAGCTGCTGCTGCCAR	188	
Qy	61	GACCTGCAGCTGGGGGCTGTCACTGTTGTTCGCTGTGGGCTGTGTGTGGGCTGTGCCAGT	120	
Db	189	GACCTGCAGCTGGGGGCTGTCACTGTTGTTCGCTGTGGGCTGTGTGTGGGCTGTGCCAGT	248	
Qy	121	GAGCTGTGCTCAACAGCCCTGTGCTGTGGGCAACTTACACAGGAGGGCAGCATGACC	180	
Db	249	GAGCTGTGCTCAACAGCCCTGTGCTGTGGGCAACTTACACAGGAGGGCAGCATGACC	308	
Qy	181	ATGCCGACGTGTACTTGTTCACATGGCAGTGGCAGGGCTGTGTGCTCAAGCCCTGTGCC	240	

Db	309	ATGCGGGA	CGGTGTA	CTTTGT	CAACAT	TGGCAG	TGGCAG	CGCTGG	GTCA	GGCCCT	GGCCCT	GGCC	368		
Qy	241	CTGTGCA	CCGTGCT	CGGCCCC	CCCGAG	CTCCCG	GTGGCG	CTGTG	GA	GTGGCG	CGCG	CGAA	300		
Db	369	CTGTGCA	CCGTGCT	CGGCCCC	CCCGAG	CTCCCG	GTGGCG	CTGTG	GA	GTGGCG	CGCG	CGAA	428		
Qy	301	GTCCAG	TGGCA	CTGCA	GATCC	CTTTCA	ATGTG	CTCTCA	CTGTG	GGCCAT	GTACT	CCAC	360		
Db	429	GTCCAG	TGGCA	CTGCA	GATCC	CTTTCA	ATGTG	CTCTCA	CTGTG	GGCCAT	GTACT	CCAC	488		
Qy	361	GCCCTG	TGA	AGCC	TCGAC	CACTAC	TGAG	CGGTG	CA	CTGCG	CGGAC	CTTAC	ATGGCCAG	420	
Db	489	GCCCTG	TGA	AGCC	TCGAC	CACTAC	TGAG	CGGTG	CA	CTGCG	CGGAC	CTTAC	ATGGCCAG	548	
Qy	421	GTGTAA	CAACA	CGCGG	CA	CGTGTG	CGGCTT	CGTGTG	GGGGT	GGCGCG	CTGTG	CA	CCAGCTTC	480	
Db	549	GTGTAA	CAACA	CGCGG	CA	CGTGTG	CGGCTT	CGTGTG	GGGGT	GGCGCG	CTGTG	CA	CCAGCTTC	608	
Qy	481	TCTCTG	CTGCT	CTTTCT	CAAT	CTGTGA	CGCA	ATGTG	TC	CA	CCCGCG	GTGA	AGTGGCCAA	540	
Db	609	TCTCTG	CTGCT	CTTTCT	CAAT	CTGTGA	CGCA	ATGTG	TC	CA	CCCGCG	GTGA	AGTGGCCAA	668	
Qy	541	ATGCA	GAACG	AGAG	GTGCG	CAACG	CGACG	CTGGT	GT	CAAT	CGGCT	CA	GTGTG	CCAGCA	600
Db	669	ATGCA	GAACG	AGAG	GTGCG	CAACG	CGACG	CTGGT	GT	CAAT	CGGCT	CA	GTGTG	CCAGCA	728
Qy	601	CTGGCA	CCCTCT	TAACG	CGCTGT	GTGTA	CTCTCC	CGCT	TCG	CA	GAGGAG	GA	CA	CGCCCTG	660
Db	729	CTGGCA	CCCTCT	TAACG	CGCTGT	GTGTA	CTCTCC	CGCT	TCG	CA	GAGGAG	GA	CA	CGCCCTG	788
Qy	661	GACCG	GA	CA	CGGCG	CGGTGA	GAC	CCCTCG	CA	CA	GAGCTG	GTGTG	CCAC	CCGTGTG	720
Db	789	GACCG	GA	CA	CGGCG	CGGTGA	GAC	CCCTCG	CA	CA	GAGCTG	GTGTG	CCAC	CCGTGTG	848
Qy	721	ACCGA	GTTTGG	GCCT	CTGGA	CGCA	CA	CTTA	CTTA	TCTG	CTGGGG	CA	CA	CGGTATATC	780
Db	849	ACCGA	GTTTGG	GCCT	CTGGA	CGCA	CA	CTTA	CTTA	TCTG	CTGGGG	CA	CA	CGGTATATC	908
Qy	781	TCGGA	GAGGA	AGCC	GTGA	CGCA	CA	CTTAC	CTGGGG	CTTAC	TGCA	AGAT	TTTC	840	
Db	909	TCGGA	GAGGA	AGCC	GTGA	CGCA	CA	CTTAC	CTGGGG	CTTAC	TGCA	AGAT	TTTC	968	
Qy	841	TCCAA	ACTCTG	GCCTT	CTC	CAGAG	AGTTT	GTGA	CA	CA	CA	CTTCTT	CA	CGCTACATGAAC	900
Db	969	TCCAA	ACTCTG	GCCTT	CTC	CAGAG	AGTTT	GTGA	CA	CA	CA	CTTCTT	CA	CGCTACATGAAC	1028
Qy	901	CAGAG	CTTCCC	CAGCA	AGCTT	CA	ACG	GTG	ATGA	AAAA	AGCTG	CCCTG	CGGGAG	ACCGGAC	960
Db	1029	CAGAG	CTTCCC	CAGCA	AGCTT	CA	ACG	GTG	ATGA	AAAA	AGCTG	CCCTG	CGGGAG	ACCGGAC	1088
Qy	961	TGCTC	CCCCG	GA	CA	CA	TGGGG	GTG	CA	CA	GATG	CTGG	CGTGA	1002	
Db	1089	TGCTC	CCCCG	GA	CA	CA	TGGGG	GTG	CA	CA	GATG	CTGG	CGTGA	1130	

RESULT 10				
AX230144				
LOCUS	AX230144	1750 bp	DNA	linear
DEFINITION	Sequence 31 from Patent WO0162797.			
ACCESSION	AX230144			
VERSION	AX230144.1	GI:15592165		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
REFERENCE	1			
AUTHORS	Vogel, G., Wood, L.S., Parodi, L.A. and Lind, P.			
TITLE	Novel G protein-coupled receptors			
JOURNAL	Patent: WO 0162797-A 31 30-AUG-2001;			
	PHARMACIA & UPJOHN COMPANY (US)			
FEATURES	Location/Qualifiers			

```
source 1..1750
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1000.4; DB 6; Length 1750;
Best Local Similarity 99.9%; Pred. No. 6,4e-163;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTGAGCTGCGAGCTGTTCACAGGCAAGGCTGTGAGAGCTGCTGTGCTCCAG 60
DB 139 ATGTGAGCTGCGAGCTGTTCACAGGCAAGGCTGTGAGAGCTGCTGTGCTCCAG 198
QY 61 GACCTGAGCTGGGGCTGTACATGTGTGCTGTGGGCTGTGGTGGGGCGGCGCAGG 120
DB 199 GACCTGAGCTGGGGCTGTACATGTGTGCTGTGGGCTGTGGTGGGGCGGCGCAGG 258
QY 121 GGGCTGTGTACACAGCCCTGTGTGTGCTGTGGCCACTACAGCAAGGCCAGCATAGC 180
DB 259 GGGCTGTGTACACAGCCCTGTGTGTGCTGTGGCCACTACAGCAAGGCCAGCATAGC 318
QY 181 ATGCCGGAAGTGTACTTTGTCAACATGAGCAATGAGCGCTGTGTCTACAGCCCTG 240
DB 319 ATGCCGGAAGTGTACTTTGTCAACATGAGCAATGAGCGCTGTGTCTACAGCCCTG 378
QY 241 CCTGTGCACTGTGCTGGCCCGGAGCTCCCGGTGGGCGCTGTGAGAGTGGGGCGGCA 300
DB 379 CCTGTGCACTGTGCTGGCCCGGAGCTCCCGGTGGGCGCTGTGAGAGTGGGGCGGCA 438
QY 301 GTCCAGTGGCACTGTGAGATCCCTTCAATGTGTCTCACTGTGTGCCATGTACTCA 360
DB 439 GTCCAGTGGCACTGTGAGATCCCTTCAATGTGTCTCACTGTGTGCCATGTACTCA 498
QY 361 GGCCTGTGAGCTGTGCAACCACTACATGAGCTGCACTCCCGGCACTTACATGCGCAG 420
DB 499 GGCCTGTGAGCTGTGCAACCACTACATGAGCTGCACTCCCGGCACTTACATGCGCAG 558
QY 421 GTGTACAACACGCGGCACTGTGGGCTTGTGTGGGGTGGGCGCTGTGACAGCTTC 480
DB 559 GTGTACAACACGCGGCACTGTGGGCTTGTGTGGGGTGGGCGCTGTGACAGCTTC 618
QY 481 TCTGTGCTGCTTTTCAATCTGTGAGCCATGTGTCCACCCGCGCTAGAGTGGCCAG 540
DB 619 TCTGTGCTGCTTTTCAATCTGTGAGCCATGTGTCCACCCGCGCTAGAGTGGCCAG 678
QY 541 ATGCAAGAAGCAAGCTGCGGCAAGCCAGCGCTGTGTATATGCTGCACTGTGTCCAG 600
DB 679 ATGCAAGAAGCAAGCTGCGGCAAGCCAGCGCTGTGTATATGCTGCACTGTGTCCAG 738
QY 601 CTGGCCACCTCTACGCGCTGTGTACTCTCCCGGTCCGCAAGGAGGACACGCCCC 660
DB 739 CTGGCCACCTCTACGCGCTGTGTACTCTCCCGGTCCGCAAGGAGGACACGCCCC 798
QY 661 GACCGGAGACGCGGCCGCTGTGAGCCCTTGTGGCACAAGCTGTGTGTGGCCACG 720
DB 799 GACCGGAGACGCGGCCGCTGTGAGCCCTTGTGGCACAAGCTGTGTGTGGCCACG 858
QY 721 AGCGAGTTTGGGCTGTGAGCGCCACATATCTGTCTGTGGGCAACAGCTATCATC 780
DB 859 AGCGAGTTTGGGCTGTGAGCGCCACATATCTGTCTGTGGGCAACAGCTATCATC 918
QY 781 TCGCAGGAGGAGCCCGTGTGACGCACTAATCTGGGGCTACTGCACTTTGTGAGGATTC 840
DB 919 TCGCAGGAGGAGCCCGTGTGACGCACTAATCTGGGGCTACTGCACTTTGTGAGGATTC 978
QY 841 TCCAAACTCTGGCTTTCTCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAG 900
DB 979 TCCAAACTCTGGCTTTCTCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAG 1038
QY 901 CAGAGCTTCCCGAGAGCTCCACGCGCTGATGAAAAAGCTGCTGTGGGGGACCGGGAC 960
DB 1039 CAGAGCTTCCCGAGAGCTCCACGCGCTGATGAAAAAGCTGCTGTGGGGGACCGGGAC 1098
```

```
QY 961 TGCCTCCCGAGACCATATGGGGGTGACAGAGGTGCTGGCGTAG 1002
DB 1099 TGCCTCCCGAGACCATATGGGGGTGACAGAGGTGCTGGCGTAG 1140
RESULT 11
BD231040
LOCUS BD231040 1878 bp DNA linear PAT 17-JUL-2003
DEFINITION Human GPCR proteins.
ACCESSION BD231040
VERSION BD231040.1 GI:33040810
KEYWORDS JP 2002525054-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1878)
Bandman, O., Lai, P., Tang, T.Y., Corley, N.C., Guegler, K.J.,
Gorgone, G.A., and Baughn, M.R.
REFERENCE
AUTHORS Human GPCR proteins
PATENT: JP 2002525054-A 3 13-AUG-2002;
JOURNAL INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002525054-A/3
PD 13-AUG-2002
PF 17-SEP-1999 JP 2000570320
PR 17-SEP-1998 US 09/156513
PI OLGA BANDMAN, PREETI LAL, TOM Y TANG, NEIL C CORLEY, KARL J PI
GUEGLER,
PI GINA A GORGONE, MARIAH R BAUSHN
PC C12N15/09, A61K38/00, A61K39/395, A61K39/395, A61K45/00, A61P1/04,
A61P1/15,
A61P1/18, A61P3/00, A61P3/10, A61P5/16, A61P7/00, A61P7/04, A61P7/06, PC
A61P9/10,
PC A61P9/10, A61P11/00, A61P11/06, A61P13/02, A61P13/12, A61P17/00, PC
A61P17/02,
PC A61P17/06, A61P17/12, A61P19/00, A61P19/02, A61P19/06, A61P19/08,
PC A61P19/10,
PC A61P21/02, A61P21/04, A61P25/00, A61P25/02, A61P25/08, A61P25/14,
PC A61P25/16,
PC A61P25/18, A61P25/28, A61P27/02, A61P29/00, A61P29/00, A61P31/00,
PC A61P31/04,
PC A61P31/10, A61P31/12, A61P31/18, A61P35/00, A61P35/02, A61P37/00,
PC A61P37/02,
PC A61P37/08, C07K14/705, C12N1/15, C12N1/19, C12N1/21, C12N5/10, PC
C12P21/02,
PC C12Q1/68, C12N15/00, C12N5/00, A61K37/02
CC Incyte ID No: 2214673CB1
FH Key source location/Qualifiers
FT source 1..1878 location/Qualifiers
FT 1..1878 /organism="Homo sapiens (human)".
FEATURES
source 1..1878
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.8%; Score 1000.4; DB 6; Length 1878;
Best Local Similarity 99.9%; Pred. No. 6,4e-163;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTGAGCTGCGAGCTGTTCACAGGCAAGGCTGTGAGAGCTGCTGTGCTCCAG 60
DB 139 ATGTGAGCTGCGAGCTGTTCACAGGCAAGGCTGTGAGAGCTGCTGTGCTCCAG 198
QY 61 GACCTGAGCTGGGGCTGTACATGTGTGCTGTGGGCTGTGGTGGGGCGGCGCAGG 120
DB 199 GACCTGAGCTGGGGCTGTACATGTGTGCTGTGGGCTGTGGTGGGGCGGCGCAGG 258
QY 121 GGGCTGTGTACACAGCCCTGTGTGTGCTGTGGCCACTACAGCAAGGCCAGCATAGC 180
```

```

Db      259 GGGCTGTGCTACAAAGCCCGTGGTGGTGGCCACTACAGCAAGGCAAGATGAC 318
Qy      181 ATGCCGAGAGTGTACTTTTGTCAACATGGCAGTGGCAGGCTGGTCTCAGCGCCCTGGGC 240
Db      319 ATGCCGAGAGTGTACTTTTGTCAACATGGCAGTGGCAGGCTGGTCTCAGCGCCCTGGGC 378
Qy      241 CCTGTGACCTGTGGTGGGCGCCGAGAGTCCCGGTGGGCGCTGTGGAGTGTGGGCGCGAA 300
Db      379 CTTGTGACCTGTGGTGGGCGCCGAGAGTCCCGGTGGGCGCTGTGGAGTGTGGGCGCGAA 438
Qy      301 GTTCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 360
Db      439 GTTCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 498
Qy      361 GCCCTGTGAGCTTGCAGCACTACATGAGCGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 420
Db      499 GCCCTGTGAGCTTGCAGCACTACATGAGCGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG 558
Qy      421 GTGTCAACACGCGGAGCGTGGGCGCTGTGGTGGGCGTGGGCGTGGTGGTGGTGGTGGTGG 480
Db      559 GTGTCAACACGCGGAGCGTGGGCGCTGTGGTGGGCGTGGGCGTGGTGGTGGTGGTGGTGG 618
Qy      481 TCTCTGCTGCTCTTCTACATCTGCAAGCAGTGTGCCAGCGCGCTAGAGTGGCGCAAG 540
Db      619 TCTCTGCTGCTCTTCTACATCTGCAAGCAGTGTGCCAGCGCGCTAGAGTGGCGCAAG 678
Qy      541 ATGCAGACGACAGAGCTGCCAGCGCAGCAGCTGTGTTCATCGGCTAGCTGTGGCAGCA 600
Db      679 ATGCAGACGACAGAGCTGCCAGCGCAGCAGCTGTGTTCATCGGCTAGCTGTGGCAGCA 738
Qy      601 CTGGCCACCTCTACAGCGGCTGTGTCTCTCCGCGTCCGAGGAGGAGCAAGCCCGTGG 660
Db      739 CTGGCCACCTCTACAGCGGCTGTGTCTCTCCGCGTCCGAGGAGGAGCAAGCCCGTGG 798
Qy      661 GACCGGAGACGCGGCGCGGCTGGAGCCCTGGGCAACAGGCTGTGGTGGCAGCCGTGGC 720
Db      799 GACCGGAGACGCGGCGCGGCTGGAGCCCTGGGCAACAGGCTGTGGTGGCAGCCGTGGC 858
Qy      721 AGCGAGTTGGGCTGTGAGCGCAGCACTATCTGATCTCTGTGGGAGCAACGCTCATCATC 780
Db      859 AGCGAGTTGGGCTGTGAGCGCAGCACTATCTGATCTCTGTGGGAGCAACGCTCATCATC 918
Qy      781 TGGCGAGGAGACCGGCTGTGAGCGCAGCACTATCTGAGGCTGTGACTTTGTGAAGATTTC 840
Db      919 TGGCGAGGAGACCGGCTGTGAGCGCAGCACTATCTGAGGCTGTGACTTTGTGAAGATTTC 978
Qy      841 TCCAAACTCTGGCTTCTCAGCAGCTTGTGACACCACTTCTTACCGCTACATGAGC 900
Db      979 TCCAAACTCTGGCTTCTCAGCAGCTTGTGACACCACTTCTTACCGCTACATGAGC 1038
Qy      901 CAGAGCTTCCCGAGCAAGCTCCAAAGGCTGATGAAAAAGCTGGCCCTGGGGAGCCGGCAG 960
Db      1039 CAGAGCTTCCCGAGCAAGCTCCAAAGGCTGATGAAAAAGCTGGCCCTGGGGAGCCGGCAG 1098
Qy      961 TGTCTCCCGAGCAACATGGGGGTGGCAGCAGTGTCTGGCGTAG 1002
Db      1099 TGTCTCCCGAGCAACATGGGGGTGGCAGCAGTGTCTGGCGTAG 1140

```

```

RESULT 12
AB093636 1998 bp mRNA 1linear PRI 17-OCT-2002
LOCUS AB093636
DEFINITION Macaca fascicularis brain cDNA clone:Ofla-17705, full insert
sequence.
ACCESSION AB093636
VERSION AB093636.1 GI:24059728
KEYWORDS oligo capping; file (full insert sequence)
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Macaca.

```

```

REFERENCE
AUTHORS Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirata, M.,
Suto, Y., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
TITLE Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
chromosomes
JOURNAL Gene 275 (1), 31-37 (2001)
MEDLINE 21458551
PUBMED 11574149
REFERENCE 2 (bases 1 to 1998)
AUTHORS Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
TITLE Direct Submision
JOURNAL Submitted (16-OCT-2002) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genbank/,
Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
COMMENT
Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CAGCATGTG)
R. Site2: DraIII (GACCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by oligo-capping method. Custom primers used for
sequencing ( 5' end primer [CTTCTCTTAAAGCTGCG] ;
3' end primer [CGACCTGACCTGACGAC] ).
FEATURES
source
1. 1998
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="Ofla-17705"
/sex="male"
/tissue_type="frontal lobe left"
/clone_lib="macaque brain cDNA library Ofla"
/dev_stage="adult"
222..1223
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC21610.1"
/db_xref="GI:24059729"
/translation="MVGCGWENGTLVBEIPACQDQLQSLSLILGLVGVPLCY
NALVLTAMSHKASMTMPDVFVNNAVAGLVLSALAPVLLGPSSOMLWAGSEVH
VALQIPFNVSLSLVNASTALSLDHYIRALPRTMASVYNTFRHGVFGVGGALLTSF
SLFLYICSHVSTRALBCAKMNAEAATLVEIGVVPALAAVLAALVLSVRBET
PLDRGTGLBPSAHLLVAVTCTGTGMLTPHVLVILGHTVLSRQPDAAHYLGILHF
VVDLSKLAFSSSFVPLVRYMNSFFPKLRMLKKPLCGDRHCSPHHTGVQOYLA"
CDS
91..441 Score 915.6; DB 9; Length 1998;
Query Match 94.6%; Pred. No. 2.7e-148;
Best Local Similarity 0; Mismatches 54; Indels 0; Gaps 0;
Matches 948; Conservative 0;
ORIGIN
1 ATGTGAGCTGACAGTGTTCACAGCAGCAGGCTGTGTGAGAGAGCTGCTCTGCGCAG 60
222 ATGTGGGGCTGGCGGCTGTTCAACGACAGGCTGTGTGAGAGAGCTGCTCTGCGCAG 281
Qy 61 GACCTGACAGTGGGCTGTCTGTTGTGCTGCTGCGGCTGTGTGAGGCTGCGCAGTGG 120
Db 282 GACCTGACAGTGGGCTGTCTGTTGTGCTGCTGCGGCTGTGTGAGGCTGCGCAGTGG 341
Qy 121 GGCCTGTGCTACAGCGCCCTGTGTGCTGCGCACTACACAGCAGGCGCAGATGACC 180
Db 342 GGCCTGTGCTACAGCGCCCTGTGTGCTGCGCACTACACAGCAGGCGCAGATGACC 401
Qy 181 ATGCCGAGCTGTACTTTGTCAACATGGCAGTGGCAGGCTGTGTCTACGCGCCCTGGCC 240
Db 402 ATGCCGAGCTGTACTTTGTCAACATGGCAGTGGCAGGCTGTGTCTACGCGCCCTGGCC 461

```

```

Qy 241 CCTGTGACCTGCTGCGCCCCCGAGCTCCCGTGGGGCGTGTGAGATGTGGCGCGGAA 300
Db 462 CCTGTGACCTGCTGCGCCCCCGAGCTCCCGTGGGGCGTGTGAGATGTGGCGCGGAG 521
Qy 301 GTTCCACGTGGGCACTGACAGATCCCTTCAATGTGTCTTCACTGTGGGCATGTACTCCACC 360
Db 522 GTTCCACGTGGGCACTGACAGATCCCTTCAATGTGTGTCTTCACTGTGGGCATGTACTCCACT 581
Qy 361 GCGCTGCTGAGCGCTGACCACTTACATGAGGTGTGACCTGCGCGGACCTTACATGCGGACG 420
Db 562 GCGCTGCTGAGCGCTGACCACTTACATGAGGTGTGACCTGCGCGGACCTTACATGCGGACG 641
Qy 421 GTTACACACAGCGGCAAGTGTGGCGGTCTGTGTGGGGTGGCGCGCTGTGACCAAGCTTC 480
Db 642 GTTACACACAGCGGCAAGTGTGGCGGTCTGTGTGGGGTGGCGCGCTGTGACCAAGCTTC 701
Qy 481 TCCCTGCTGCTCTTCTTACATGTGTGACCACTGTGTGACCCGCGCGCTAGAGTGGCGGAA 540
Db 702 TCCCTGCTGCTCTTCTTACATGTGTGACCACTGTGTGACCCGCGCGCTAGAGTGGCGGAA 761
Qy 541 ATGACGAAACGAGAAAGCTGCGGACGCGGCTGTGTGACCACTGTGTGACCACTGTGTGAC 600
Db 762 ATGACGAAACGAGAAAGCTGCGGACGCGGCTGTGTGACCACTGTGTGACCACTGTGTGAC 821
Qy 601 CTGCGCACCTCTTACAGCGGTGTGTGACTCTCTCCGCGTCCGACGAGGAGACAGCCCTTG 660
Db 822 CTGCGCACCTCTTACAGCGGTGTGTGACTCTCTCCGCGTCCGACGAGGAGACACACCCCTTG 881
Qy 661 GACCGGAGACAGCGGCGCGGTGTGAGCCCTGTGGGACACAGGCTGTGTGACCACTGTGTGAC 720
Db 882 GACCGGAGACAGCGGCGCGGTGTGAGCCCTGTGGGACACAGGCTGTGTGACCACTGTGTGAC 941
Qy 721 AGCGAGTTTGGGCTGTGAGCGCCACACTATCTGTGCTGTGGGGACACGGTCACTATC 780
Db 942 ACCGAGTTTGGGCTGTGAGCGCCACACTATCTGTGCTGTGGGGACACAGTCTCATC 1001
Qy 781 TCGCGAGGAGAAAGCGCGTGTGACGACACTTCTGCGGCTTACTGCTCTTGTGAAAGATTTC 840
Db 1002 TCGCGAGGAGAAAGCGCGTGTGACGACACTTCTGCGGCTTACTGCTCTTGTGAAAGATTTC 1061
Qy 841 TCGAACTCTGCTGCTTCTTCCAGAGCTTGTGACACCACTTCTTACCGCTTACATGTAGAC 900
Db 1062 TCGAACTCTGCTGCTTCTTCCAGAGCTTGTGACACCACTTCTTACCGCTTACATGTAGAC 1121
Qy 901 CAGAGCTTCCCAAGCAAGCTTCAACGGCTGTGATAAAGCTGCTGTGCGGGACCGGACG 960
Db 1122 CAGAGCTTCCCAAGCAAGCTTCCGCGGGCTGTGATAAAGCTGCTGTGCGGGACCGGACG 1181
Qy 961 TCGTCCCCCGAGCAATGGGGGTGTGACGAGGCTGTGGCTAG 1002
Db 1182 TCGTCCCCCGAGCAATGGGGGTGTGACGAGGCTGTGGCTAG 1223

```

```

RESULT 13
AK097629/c 2240 bp mRNA linear PRI 30-JAN-2004
LOCUS AK097629 Homo sapiens cDNA FLJ40310 f18, clone TEST12029672.
DEFINITION AK097629
ACCESSION AK097629
VERSION AK097629.1 GI:21757459
KEYWORDS oligo capping; f18 (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE

```

AUTHORS
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obase, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,

```

```

Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Houchi, T., Kusano, Y., Kanehori, K., Takahashi, Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takesuchi, K., Arita, M., Imose, N., Mueshishima, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Teraishi, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, T.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, K., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuo, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039
2

```

JOURNAL

PUBMED

REFERENCES

AUTHORS

TITLE

```

TITLE
JOURNAL
PUBMED
REFERENCES
AUTHORS

```

JOURNAL

PUBMED

REFERENCES

AUTHORS

COMMENT

```

COMMENT
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2240
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST12029672"
/tissue_type="testis"
/clone_lib="TEST12"
/note="Cloning vector: PME18SFu3"

```

FEATURES

source

ORIGIN

```

Query Match 87.6%; Score 877.8; DB 9; Length 2240;
Best Local Similarity 99.8%; Pred. NO. 9e-142;
Matches 879; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 122 GCTGTGCTACACCGCTGTGTGTGCTGCGCACTTACACAGAGCCGACATGACCA 181
Db 1977 GCTGTGTCTACACGCGCTGTGTGTGCTGCGCACTTACACAGAGCCGACATGACCA 1918
Qy 182 TGCCGAGAGTGTACTTGTGTCAACATGGCAGTGTGAGGCTGTGTCTAGGCGCCCTGCGCC 241
Db 1917 TGCCGAGAGTGTACTTGTGTCAACATGGCAGTGTGAGGCTGTGTCTAGGCGCCCTGCGCC 1858

```

```

Oy 242 CTGTGACCTGCTGGGCCCCCGAGCTCCCGGCGGCGCTGTGTGAGATGTGGCGGGGANG 301
Db 1857 CCGTGACCTGCTGGGCCCCCGAGCTCCCGGCGGCGCTGTGTGAGATGTGGCGGGGANG 1798
Oy 302 TCCACGTGGACATGACAGATCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCCACGG 361
Db 1797 TCCAGTGGACATGACAGATCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCCACGG 1738
Oy 362 CCGTGTGAGCTTGCACATCATGTAGAGGTGACTGCCCGGACCTTACATGCGCAAG 421
Db 1737 CCGTGTGAGCTTGCACATCATGTAGAGGTGACTGCCCGGACCTTACATGCGCAAG 1678
Oy 422 TGTACAAACGCGGCGCGTGTGGCTGTGTGGGTGGCGCGCTGTGACCAAGTTCT 481
Db 1677 TGTACAAACGCGGCGCGTGTGGCTGTGTGGGTGGCGCGCTGTGACCAAGTTCT 1618
Oy 482 CCGTGTGCTTCTTCACTGTGAGCCATGTGTCCACCGCGCGCTAGAGTGGCCAGA 541
Db 1617 CCGTGTGCTTCTTCACTGTGAGCCATGTGTCCACCGCGCGCTAGAGTGGCCAGA 1558
Oy 542 TGCAGAACGCAAGAGTGTCCGACCGACGCTGTGTTCATCGGCTACGTGTGCCAGCAC 601
Db 1557 TGCAGAACGCAAGAGTGTCCGACCGACGCTGTGTTCATCGGCTACGTGTGCCAGCAC 1498
Oy 602 TGGCCACCTCTTACCGGCGTGTGTCTCTCCCGCGCTCCGACGAGGACACCGCCCTGG 661
Db 1497 TGGCCACCTCTTACCGGCGTGTGTCTCTCCCGCGCTCCGACGAGGACACCGCCCTGG 1438
Oy 662 ACCGGGACACGGGCGCGCTGTGAGCCCTCGGACACAGAGCTGTGTGGCCACCGTGTGCA 721
Db 1437 ACCGGGACACGGGCGCGCTGTGAGCCCTCGGACACAGAGCTGTGTGGCCACCGTGTGCA 1378
Oy 722 CGCAGTTTGGGCTGTGAGCGCCACACTATCTGTCTGTGGGACACAGGTCATCATCT 781
Db 1377 CGCAGTTTGGGCTGTGAGCGCCACACTATCTGTCTGTGGGACACAGGTCATCATCT 1318
Oy 782 CGCGAGGAGGACCGCGTGTGAGCGCCACTACTGTGGGCTACTGCACTTTGGAAGATTTCT 841
Db 1317 CGCGAGGAGGACCGCGTGTGAGCGCCACTACTGTGGGCTACTGCACTTTGGAAGATTTCT 1258
Oy 842 CCAAACTCTGGGCTTCTCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAACC 901
Db 1257 CCAAACTCTGGGCTTCTCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAACC 1198
Oy 902 AAGAGTCTCCCAAGCAAGTCTCAACGCGCTGATGAAAAGCTGCTCGGGGACCGGCACT 961
Db 1197 AAGAGTCTCCCAAGCAAGTCTCAACGCGCTGATGAAAAGCTGCTCGGGGACCGGCACT 1138
Oy 962 GCTCCCGGACCAATGGGGGTGCACAGGTGTGGCTTNG 1002
Db 1137 GCTCCCGGACCAATGGGGGTGCACAGGTGTGGCTTNG 1097

```

```

RESULT 14
BC003323 1468 bp mRNA linear ROD 29-JUN-2004
LOCUS Mus musculus cDNA sequence BC003323, mRNA (cDNA clone MGC:7035
DEFINITION IMAGE:3156060), complete cds.
ACCESSION BC003323
VERSION BC003323.1 GI:13097086
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1468)
Struhsberg, R.L., Pelngold, E.A., Grouse, L.H., Dege, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Schmeem, C.M., Schlier, G.D.,
Altekar, S.F., Zeeberg, B., Buecaw, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F.,
Diachenko, L., Marcusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Udén, T.B., Toshiyuki, S.,

```

REMARK

COMMENT

JOURNAL

TITLE

PUBMED

AUTHORS

JOURNAL

JOURNAL

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Petere, G.J.,
 Abramson, R.D., Mullaney, S.J., Bosak, S.A., McMan, P.J.,
 McKernan, K.J., Malek, J.A., Guneratne, P.H., Richard, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Weng, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smal, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

NIH-MGC Project URL: <http://mgi.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 6 Row: m Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13385047.

FEATURES

source

gene

CDS

```

1..1468
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:7035 IMAGE:3156060"
/tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_1lb="NCI CGAP_Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..1468
/gene="BC003323"
/notes="synonym: MGC7035"
/db_xref="LocustID:80290"
/db_xref="MGI:1933113"
72..1073
/gene="BC003323"
/codon_start=1
/product="cDNA sequence BC003323"
/protein_id="AAH03323.1"
/db_xref="GI:13097087"
/db_xref="LocustID:80290"
/db_xref="MGI:1933113"
/translation="MMSGCPPLTNMAEPRICRNRIGLMTSLYLNGVPSGYN
ALLVIANLASKRTMTMPDYFVNVAVGLVTLALPAVLIGFHSRMALMSSEAHV
TLILFNVASVIRIATLSTLSDYIERALPRTYASVYNTBHVCGFVGAVLTSFS
SLPLFNVASVIRIATLSTLSDYIERALPRTYASVYNTBHVCGFVGAVLTSFS
LIDPISRLDPSVHRLVAVTCVGFGLMPYVLSGLVTVLSRGTVEGHVGLDQVAK
DLAKLAPSSSVTPLVLYRIKAPFKLRRLMKHCGRRHCSPPSISIQVMAQA"

```

ORIGIN

```

Query Match      60.9%; Score 610; DB 10; Length 1468;
Best Local Similarity 76.9%; Pred. No. 1.5e-95;
Matches 771; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

OY 1 ATGTGGAGCTGTGAGCTGTTCACAGGACAGGAGCTGTGAGAGAGCTGCTCCGACG 60
DB 72 ATGTGGAGCTGTGAGCTGTTCACAGGACAGGAGCTGTGAGAGAGCTGCTCCGACG 128
OY 61 GACCTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 120
DB 129 AACCTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 188
OY 121 GACCTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 180
DB 189 AGCTTATGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 248
OY 181 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 240
DB 249 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 308
OY 241 CTTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 300
DB 309 CTTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 368
OY 301 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 360
DB 369 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 428
OY 361 GACCTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 420
DB 429 GACCTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 488
OY 421 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 480
DB 489 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 548
OY 481 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 540
DB 549 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 608
OY 541 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 600
DB 609 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 668
OY 601 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 660
DB 669 GTTCAAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 728
OY 661 GACCTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 720
DB 729 GACCTGTGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 788
OY 721 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 780
DB 789 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 845
OY 781 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 840
DB 846 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 905
OY 841 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 900
DB 906 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 965
OY 901 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 960
DB 966 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 1025
OY 961 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 1025
DB 1026 ATGCGGAGCTGTGAGCTGTTCACAGGAGCTGTGAGAGAGCTGCTCCGACG 1067

```

```

RESULT 15
AC140216
LOCUS
DEFINITION
AC140216 194560 bp DNA linear ROD 10-JUN-2004
Mus musculus BAC clone RP23-276B20 from chromosome 5, complete
sequence.
AC140216.3 GI:48675535
HTG.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 194560)
Bergthoff, A., Cotton, M., Bielicki, L. and Haglund, K.
The sequence of Mus musculus BAC clone RP23-276B20
Unpublished (2001)
2 (bases 1 to 194560)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 194560)
Wilson, R.K.
Direct Submission
Submitted (08-MAY-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 194560)
Wilson, R.K.
Direct Submission
Submitted (12-JUN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 194560)
Wilson, R.K.
Direct Submission
Submitted (10-JUL-2004) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 12, 2004 this sequence version replaced gi:47084696.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@genome.wustl.edu
----- Summary Statistics
Center project name: M_BA0276B20
-----

```

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-23 BAC library has been constructed by Kazuhiro Osegawa
 and Minko Tateo in the laboratory of Pieter de Jong
 (http://www.chori.org) from female C57BL/6J mouse kidney and/or
 brain genomic DNA. The clone and detailed information can be
 obtained from Research Genetics, Inc. (http://www.resgen.com) or
 Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC144902.

FEATURES
source Location/Qualifiers

```

1..194560
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="5"
    /map="5"
    /clone="RP23-276B20"
    /clone_11b="RPC1-23"
    388..741
repeat_region /rpt_family="MaLR"
    742..1371
repeat_region /rpt_family="MaLR"
    1513..1692
repeat_region /rpt_family="B2"
    1736..1795
repeat_region /rpt_family="MaLR"
    1861..2178
repeat_region /rpt_family="MaLR"
    2181..2410
repeat_region /rpt_family="MaLR"
    4755..5099
unsure
    /note="Unresolved simple sequence repeat."
repeat_region /rpt_family="Alu"
    6703..6811
repeat_region /rpt_family="B4"
    7493..7637
repeat_region /rpt_family="B4"
    7724..7845
repeat_region /rpt_family="B4"
    7973..8026
repeat_region /rpt_family="ZP3AR_MM"
    8027..8108
repeat_region /rpt_family="ID"
    8109..8130
repeat_region /rpt_family="ZP3AR_MM"
    8131..8213
repeat_region /rpt_family="ID"
    8214..8272
repeat_region /rpt_family="ZP3AR_MM"
    8312..8591
repeat_region /rpt_family="B4"
    8879..9004
repeat_region /rpt_family="Alu"
    9367..9461
repeat_region /rpt_family="Alu"
    9469..9605
repeat_region /rpt_family="Alu"
    9926..9951
repeat_region /rpt_family="ERV1"
    9952..10045
repeat_region /rpt_family="Alu"
    10046..10107
repeat_region /rpt_family="ERV1"
    10109..10251
repeat_region /rpt_family="Alu"
    10254..10387
repeat_region /rpt_family="Alu"
    10418..10520
repeat_region /rpt_family="B4"
    10521..10650
repeat_region /rpt_family="Alu"
    10701..10857
repeat_region /rpt_family="B4"
    11131..11188
repeat_region /rpt_family="ZP3AR_MM"
    11189..11272
repeat_region /rpt_family="ID"
    11273..11294
repeat_region /rpt_family="ZP3AR_MM"
    11295..11377

```

```

repeat_region /rpt_family="ID"
    11378..11399
repeat_region /rpt_family="ZP3AR_MM"
    11400..11482
repeat_region /rpt_family="ID"
    11483..11504
repeat_region /rpt_family="ZP3AR_MM"
    11505..11587
repeat_region /rpt_family="ID"
    11588..11639
repeat_region /rpt_family="ZP3AR_MM"
    12030..12127
repeat_region /rpt_family="L1"
    12866..12934
repeat_region /rpt_family="ID"
    12946..13155
repeat_region /rpt_family="MER1_type"
    13156..13229
repeat_region /rpt_family="ZP3AR_MM"
    13231..13309
repeat_region /rpt_family="ID"
    13326..13448
repeat_region /rpt_family="B4"
    13514..13721
repeat_region /rpt_family="B2"
    13747..13903
repeat_region /rpt_family="B4"
    13904..14032
repeat_region /rpt_family="L1"
    14089..14196
repeat_region /rpt_family="B4"
    14372..14515
repeat_region /rpt_family="B4"
    14920..15033
repeat_region /rpt_family="MaLR"
    15274..15420
repeat_region /rpt_family="Alu"
    15610..15664
repeat_region /rpt_family="B2"
    15696..15746
repeat_region /rpt_family="Alu"
    15749..15914
repeat_region /rpt_family="L1"
    15915..16104
repeat_region /rpt_family="B2"
    16105..16252
repeat_region /rpt_family="L1"
    16261..16535
repeat_region /rpt_family="B4"
    16609..16737
repeat_region /rpt_family="L1"
    16983..17245
repeat_region /rpt_family="B4"
    17302..17447
repeat_region /rpt_family="Alu"
    17475..17603
repeat_region /rpt_family="Alu"
    17992..18051

```

Query Match 60.9%; Score 610; DB 10; Length 194560;
Best Local Similarity 76.9%; Pred. No. 8 6e-96;
Matches 771; Conservative 0; Mismatches 225; Indels 6; Gaps 2;

```

QY 1 ATGTGAGCTGACGTGTTCAACGACACAGGCTGTGAGAGCTGCTGCTGCAG 60
    |||||
DB 97303 ATGTGAGCTGTGGCCACTCAACAGACAGCGTGGGCTGAGAGCCGC---TGTGCCGG 97359
    |||||
QY 61 GACCTGACAGCTGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGGGCTGCCAGT 120
    |||||
DB 97360 AACCTGGGCTGGGCTGTGGCTCTCTGCTGCTTCACTTGGGAGGAGGCTCTGTG 97419
    |||||
QY 121 GGCTGTGCTACAGCCCTGCTGTGCTGTCGCAACTCAACAGGAGCCAGCATGACC 180
    |||||

```

Db 97420 AGCTTAGGCTATATATGCTCTTCTGTGCTGGCCAACTGGCCAGAAAGAACCATGACC 97479
QY 181 ATCCGAGAGTGTACTTTGTGAACATGGCAATGGAGAGGCTGTGTCTCAGCGCCCTGGCC 240
Db 97480 ATGCCGAGGTATCTTGTGAACATGGCCGTGGGGGTGTGTCTCAGCGCACTGGCA 97539
QY 241 CCTGTGCACTGTGTGGCCCCCGAGCTCCCGGTGGGGCTGTGAGTGTGGGGGGA 300
Db 97540 CCTGTGTACTGTGTGGCCCTGGCCCACTCCAGGTGGGCCCTGTGAGCTTCAGCACTGAG 97599
QY 301 GTCCAGTGGGCACTGGCAATCCCTTCAATGTGTCTCATCTGTGGCCATGTACTCCACC 360
Db 97600 GCCATGTGACATCTGTCTATCTGTCAACGTGGCTTCCCTGTGACCATGTACTCCACT 97659
QY 361 GCCCTGTGAGCTTGAACCACTAATGAGAGGTGCACTGCGCGGGAACCTTACATGGCCAGC 420
Db 97660 GCACCTGTGAGCTTGAATTAATGAGAGGTGCTGCGCAGCACTTACATGAGCCAGT 97719
QY 421 GTGTAAACAAGCGGGCACTGTGTGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 480
Db 97720 GTGTAAACAACCGGCACTGTGTGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 97779
QY 481 TCCCTGTGTCTCTTCTACATCTGTGCAATGTGTCCACCGCGGCTAAGTGTGCGCAAG 540
Db 97780 TCCCTGTGTCTCTTCTACATCTGTGCAATGTGTCCACCGCGGCTAAGTGTGCGCG 97839
QY 541 ATGTGAACCGGAAGCTGCGCAAGCGCACTGTGTTCATCGGCTACGTGTGCCAGCA 600
Db 97840 ATGTGAACCAAGGAGGCAAGCGCAATGTGTTCATCGGCTACGTGTGCCAGGT 97899
QY 601 CTGGCCACCTCTTACGCGGTGTGTACTCTCCCGCTCCGCAAGGAGGACACGCCCCCTG 660
Db 97900 CTGGCTGT 97959
QY 661 GACCGGAGACAGGCGCGGCTGTGAGCCCTGCGCACACAGGCTGTGTGTGTGTGTGTGTGT 720
Db 97960 GACCAAGACACAGGAGGCTGTGAGCCCTGCGGTGCAAGGCTGTGTGTGTGTGTGTGT 98019
QY 721 ACCGAGTTGGGCTGTGGAAGCCCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 98020 ACTGAGTTGGCTGTGGAACCTTACTTGTAGC--CTGGGGCACACAGTGTGTAGC 98076
QY 781 TCGCGAGGAGAGCCGCTGTGAGCACACTACCTGTGGGCTACTGCACTTGTGAAGATTTC 840
Db 98077 TCACGGGGGAGGACCGTGTGAGGGGCACTTATCTGGGCACTCTACAGGTGTCTAAGGACCTG 98136
QY 841 TCGAAACTCTGTGCTTCTCAAGCACTTGTGTGACACCACTTCTACCGCTACATGAAC 900
Db 98137 GCCAAGTCTGTGCTTCTCAAGCACTTGTGTGACACCACTGCTTACCGTTACATCAAC 98196
QY 901 CAGAGCTTCCCGAGCAAGCTCCACAGGCTGTGAATAAAGCTGCGCTGTGGGGAGCCGGCAC 960
Db 98197 AAAAGCTTCCCGGCAAGCTCCGGGCTGTGAATAAAGATGTCACTGTGGGGCGCGGCAC 98256
QY 961 TGTCTCCCGGAGCACATGTGGGGGTGTGAGAGGTGTGTGCGGTAG 1002
Db 98257 TGTCTCCCGGAGCCCTCGGGGATACAGAGGTGTGTGAGCAG 98298

Search completed: December 15, 2004, 00:42:49
Job time : 3379.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: December 14, 2004, 21:00:04 ; Search time 469.273 Seconds
(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225B-1

Perfect score: 1002
Sequence: 1 atgttgagagctgcagctcgtt.....tcgacgagctgcctgcgctag 1002

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1002	100.0	1002	6	ABT04866 Human G P
2	1002	100.0	1002	6	ABK6286 Human TGR
3	1002	100.0	1402	10	ADC8562 Human GPC
4	1002	100.0	1778	4	AA542859 Human G P
5	1002	100.0	1785	8	AB242598 Human G P
6	1002	100.0	1785	12	ADC30022 Human GPC
7	1002	100.0	1926	8	ACCT0841 Human G-P
8	1002	100.0	1928	8	ACA61176 Human G-P
9	1000.4	99.9	1614	3	AA290175 Human che
10	1000.4	99.8	1750	4	AA542836 Human G P
11	1000.4	99.8	1878	6	ABX73053 Human G-P
12	1000.4	99.8	1893	4	AAH99600 Human G-P
13	999	99.7	2216	8	ACCT0860 Insect CD
14	999	99.7	2258	8	ACCT0859 Insect CD
15	999	99.7	2363	3	ACCT0861 Insect CD
16	989.4	98.7	1879	3	AA290523 Human GPC
17	610	60.9	1468	12	AD030312 Human GPC
18	577	57.6	2651	4	ABK42895 Genomic s
19	577	57.6	2651	9	ADB61051 Connectiv
20	521	52.0	1269	6	ABA04303 Human ops
21	422.2	42.1	559	6	ABN26732 Human ORF

22	358.6	35.8	398	9	ACH20905
23	350.6	35.0	461	9	ACH43678
24	343.8	34.3	607	4	AA182081
25	304	30.3	497	6	ABX73103
26	275	27.4	275	6	ABX73076
27	236.4	23.6	238	6	ABX73074
28	229.4	22.9	300	3	AA290174
29	178	17.8	930	4	ABK42894
30	169	16.9	930	9	ADB61050
31	159	15.9	294	6	ABX73105
32	166	16.6	259	6	ABX73075
33	162.6	16.2	266	6	ABX73104
34	134	13.4	255	6	ABX73071
35	132	12.2	281	6	ABX73073
36	89.4	8.9	363	6	ABX73072
37	71	7.1	2932	6	AA026878
38	69.8	7.0	1089	3	AAA30581
39	69.8	7.0	1089	6	AA026832
40	69.8	7.0	1089	8	ACCT8113
41	69.8	7.0	1089	10	ADC22522
42	69.8	7.0	1089	10	ADH13995
43	69.8	7.0	1089	12	AD029766
44	69.8	7.0	1365	6	AB234905
45	69.8	7.0	1365	10	ADB84894

ALIGNMENTS

RESULT 1	ABT04866	ABT04866 standard; cDNA; 1002 BP.
ID	ABT04866	
AC	ABT04866	
XX		
DT	11-OCT-2002 (first entry)	
XX		
DE	Human G protein coupled receptor hrup28 coding sequence.	
XX		
KW	Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31;	
KW	hrup32; hrup33; hrup34; hrup35; hrup36; hrup37; gene; ss.	
OS	Homo sapiens.	
XX		
PN	WO200242461-A2.	
XX		
PD	30-MAY-2002.	
XX		
PF	26-NOV-2001; 2001WO-US044386.	
XX		
PR	27-NOV-2000; 2000US-0253404P.	
PR	12-DEC-2000; 2000US-0253466P.	
PR	20-FEB-2001; 2001US-0270266P.	
PR	20-FEB-2001; 2001US-0270286P.	
PR	06-APR-2001; 2001US-0282032P.	
PR	06-APR-2001; 2001US-0282356P.	
PR	06-APR-2001; 2001US-0282358P.	
PR	06-APR-2001; 2001US-0282359P.	
PR	14-MAY-2001; 2001US-0290917P.	
PR	31-JUL-2001; 2001US-0309208P.	
XX		
PA	(AREN-) ARENA PHARM INC.	
XX		
PI	Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;	
XX		
DR	WPI; 2002-566565/60.	
DR	P-PSDB; ABJ04068.	
PT	Novel endogenous and non-endogenous versions of G protein-coupled	
PT	receptor useful for identification of candidate compounds as receptor	
XX	agonists or antagonists for use as therapeutic agents.	
PS	Claim 3; Page 54-55; 84pp; English.	

Ach20905	Human adu
Ach3678	Human po1
AA182081	Human po1
ABX73103	Mouse var
ABX73076	Human GPC
ABX73074	Human GPC
AA290174	Fragment
ABK42894	Genomic s
ADB61050	Connectiv
ABX73105	Rat varia
ABX73075	Human GPC
ABX73104	Rat varia
ABX73071	Human GPC
ABX73073	Human GPC
ABX73072	Human GPC
AA026878	Human G-P
AAA30581	Human G-P
AA026832	Human G-P
ACCT8113	Human G-P
ADC22522	Human G-P
ADH13995	Human GPR
AD029766	Human nov
AB234905	Human gen
ADB84894	Fatnesyl

XX The present invention provides the protein and coding sequences of
 CC several human G-protein coupled receptors (GPCRs). These can be used in
 CC the identification of candidate compounds as receptor agonists or inverse
 CC agonists having applicability as therapeutic agents. The present sequence
 CC is a GPCR coding sequence of the invention

XX Sequence 1002 BP; 163 A; 333 C; 313 G; 193 T; 0 U; 0 Other;

Query Match 100.0%; Score 1002; DB 6; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 3,5e-194;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGAGCTGAGCTGTTCAACGAGCAGAGGCTGTGAGAGAGTGCCTGCTGCCAG 60
 DB 1 ATGTGAGAGCTGAGCTGTTCAACGAGCAGAGGCTGTGAGAGAGTGCCTGCTGCCAG 60
 QY 61 GACCTGCACTGGGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGGCGTGCACATG 120
 DB 61 GACCTGCACTGGGGGCTGTCACTGTTGTGCTGCTGGGCTGTGTGGCGTGCACATG 120
 QY 121 GGGCTGTGTACAAAGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 DB 121 GGGCTGTGTACAAAGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 QY 121 GGGCTGTGTACAAAGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 DB 121 GGGCTGTGTACAAAGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
 QY 181 ATGCCGAGCTGTACTTTGTCAACATGTGAGTGTGAGGCTGTGTGTGTGTGTGTGTGT 240
 DB 181 ATGCCGAGCTGTACTTTGTCAACATGTGAGTGTGAGGCTGTGTGTGTGTGTGTGTGT 240
 QY 241 CCTGTGTACCTGT 300
 DB 241 CCTGTGTACCTGT 300
 QY 301 GTTCCAGT 360
 DB 301 GTTCCAGT 360
 QY 361 GGCCTGT 420
 DB 361 GGCCTGT 420
 QY 421 GTGTACAAAGCCGCTGT 480
 DB 421 GTGTACAAAGCCGCTGT 480
 QY 481 TCCCTGT 540
 DB 481 TCCCTGT 540
 QY 541 ATGCAAGAGCAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 541 ATGCAAGAGCAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 QY 601 CTGT 660
 DB 601 CTGT 660
 QY 661 GACCGGAGCAGCGGGCTGT 720
 DB 661 GACCGGAGCAGCGGGCTGT 720
 QY 721 ACGCAGTTTGGGCTGT 780
 DB 721 ACGCAGTTTGGGCTGT 780
 QY 781 TCGGAGGAGGAGCCGCTGT 840
 DB 781 TCGGAGGAGGAGCCGCTGT 840
 QY 841 TCCAAACTCTGT 900
 DB 841 TCCAAACTCTGT 900

QY 901 CAGAGCTTCCCGAGAGCTCCAAACGCTGTGAAGTGTGCTGTGGGGGACCGGGAC 960
 DB 901 CAGAGCTTCCCGAGAGCTCCAAACGCTGTGAAGTGTGCTGTGGGGGACCGGGAC 960
 QY 961 TGTCTCCCGGAGCAGCATGTGGGGGTGCAGCAGGTGTGTGTGTGTGTGTGTGTGT 1002
 DB 961 TGTCTCCCGGAGCAGCATGTGGGGGTGCAGCAGGTGTGTGTGTGTGTGTGTGTGT 1002

RESULT 2
 ABRK6286
 ID ABRK6286 standard; cDNA; 1002 BP.

AC ABRK6286;
 DT 27-AUG-2002 (first entry)

DE Human TGR339 cDNA.

XX Human; TGR339; gene; ser; G-protein coupled receptor; GPCR; TGR;
 KW TGR-associated disorder; signal transduction; renal failure; nephritis;
 KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;
 KW diabetes insipidus; hyperprolactinemia; thirst disturbance; appetite;
 KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;
 KW circadian rhythm.

OS Homo sapiens.

Key Location/Qualifiers
 FH 1..1002
 FT /*tag= a
 FT /product= "Human TGR339 protein"

PN MO200242458-A2.

PD 30-MAY-2002.

PF 21-NOV-2001; 2001WO-US043404.

PR 22-NOV-2000; 2000US-0252841P.

PR 22-DEC-2000; 2000US-0257636P.

PR 12-JAN-2001; 2001US-0261377P.

PR 28-MAR-2001; 2001US-0279554P.

PR 29-MAR-2001; 2001US-0280696P.

XX (TTLA-) TULARIX INC.

XX Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupta JS;

XX WPI; 2002-463633/49.

XX P-PSDB; MAU97734.

PT New isolated G-protein couple receptor polypeptide, termed TGR, for

PT diagnosis and treatment of diseases such as renal failure, nephritis,

PT hypothyroidism, diabetes insipidus, and disturbances of thirst and sleep.

XX Claim 42; Page 73; 98pp; English.

XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),
 CC termed TGR, and its associated nucleic acid. The sequences of the
 CC invention are useful for identifying a compound that modulates signal
 CC transduction and for identifying a mammal having a TGR-associated
 CC disorder. The proteins and nucleic acids are useful in diagnosis and
 CC treatment of diseases or conditions such as renal failure, nephritis,
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,
 CC diabetes insipidus, hyperprolactinemia and disturbances of thirst,
 CC sleep, temperature regulation, appetite, blood pressure or any other
 CC syndrome or disease associated with the hypothalamus. The sequences can
 CC be used in regulation of circadian rhythms, for use as genetic markers
 CC for the identification of mutations associated with diseases resulting
 CC from GPCR inactivation in particular cell types and for identification of
 CC modulators of GPCR signal transduction. This sequence represents cDNA
 CC encoding the human TGR339 polypeptide

XX Sequence 1002 BP; 163 A; 333 C; 313 G; 193 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 1002; DB 6; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 3.5e-194;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGTGAGCTGACGCTGTTCAACGGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 60
DB 1 ATGTGAGCTGACGCTGTTCAACGGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 60
QY 61 GACCTGACAGCTGGGCTGTCACTGTGTGCTGTGAGGCTGTGTGAGGCTGTGCCAGT 120
DB 61 GACCTGACAGCTGGGCTGTCACTGTGTGCTGTGAGGCTGTGTGAGGCTGTGCCAGT 120
QY 121 GGCCTGTGTCTAACAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 121 GGCCTGTGTCTAACAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 181 ATGCCGGAGCTGTACTTTGTCAACATGTGAGTGTGAGGCTGTGTGTGTGTGTGTGT 240
DB 181 ATGCCGGAGCTGTACTTTGTCAACATGTGAGTGTGAGGCTGTGTGTGTGTGTGTGT 240
QY 241 CCTGTGACAGCTGTGAGCCCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 300
DB 241 CCTGTGACAGCTGTGAGCCCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 300
QY 301 GTTCAAGTGTGACAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 360
DB 301 GTTCAAGTGTGACAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 360
QY 361 GGCCTGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 420
DB 361 GGCCTGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 420
QY 421 GTGTACACACAGCGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 421 GTGTACACACAGCGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 481 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
DB 481 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 541 ATGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 600
DB 541 ATGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 600
QY 601 CTGGCCACCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 660
DB 601 CTGGCCACCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 660
QY 661 GACCGGAGACAGCGGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 720
DB 661 GACCGGAGACAGCGGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 720
QY 721 AGCGAGTTTGGGCTGTGAGCGCCACTATCTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 721 AGCGAGTTTGGGCTGTGAGCGCCACTATCTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 781 TCGCGAGGAGAGCCGCTGTGAGCGCACTACTGTGAGGCTGTGAGCTGTGTGTGTGT 840
DB 781 TCGCGAGGAGAGCCGCTGTGAGCGCACTACTGTGAGGCTGTGAGCTGTGTGTGTGT 840
QY 841 TCCAAATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 841 TCCAAATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 901 CAGAGCTTCCCGACAGAGCTTCCACGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 901 CAGAGCTTCCCGACAGAGCTTCCACGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
QY 961 TGTCTCCCGACACATGTGGGGGTGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 1002
  
```

DB 961 TGTCTCCCGACACATGTGGGGGTGAGAGGTGTGTGTGTGTGTGTGTGTGTGT 1002

RESULT 3
 AD86562
 ID AD86562 standard; DNA; 1402 BP.

AD86562;
 01-JAN-2004 (first entry)

Human GPCR gene SEQ ID NO:1015.

de; gene: human; GPCR;
 guanosine triphosphate-binding protein coupled receptor; gene therapy.

Homo sapiens.

EP1270724-A2.

02-JAN-2003.

18-JUN-2002; 2002EP-00013517.

18-JUN-2001; 2001JP-00246789.

(NAT INST ADVANCED IND SCI & TECHNOLOGY.
 (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATION.

Suwa M, Arai K, Akiyama Y, Aburatani H;

WPI: 2003-315783/31.

P-PsDB; AD86563.

New polynucleotide, useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor.

Claim 1, SEQ ID NO 1015; 28pp; English.

The invention relates to a novel polynucleotide encoding a guanosine CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of CC the invention may have a use in gene therapy. The polynucleotide and CC polypeptide are useful for preparing a composition for treating a patient CC in need of increased or suppressed activity or expression of the CC guanosine triphosphate-binding protein coupled receptor. The CC polynucleotide sequences shown in AD85548-AD87616 encode GPCR's of the CC invention.

Sequence 1402 BP; 238 A; 456 C; 428 G; 280 T; 0 U; 0 Other;

Query Match 100.0%; Score 1002; DB 10; Length 1402;
 Best Local Similarity 100.0%; Pred. No. 3.7e-194;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGTGAGCTGACGCTGTTCAACGGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 60
DB 201 ATGTGAGCTGACGCTGTTCAACGGACAGGGCTGTGTGAGAGAGCTGCTGCTGCCAG 260
QY 61 GACCTGACAGCTGGGCTGTCACTGTGTGCTGTGAGGCTGTGTGAGGCTGTGCCAGT 120
DB 261 GACCTGACAGCTGGGCTGTCACTGTGTGCTGTGAGGCTGTGTGAGGCTGTGCCAGT 120
QY 121 GGCCTGTGTCTAACAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 321 GGCCTGTGTCTAACAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 181 ATGCCGGAGCTGTACTTTGTCAACATGTGAGTGTGAGGCTGTGTGTGTGTGTGTGT 240
DB 381 ATGCCGGAGCTGTACTTTGTCAACATGTGAGTGTGAGGCTGTGTGTGTGTGTGTGT 240
QY 241 CCTGTGACAGCTGTGAGCCCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 300
  
```

```

Db      441  CCGTGACACTGCTCGGCCCCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAA 500
Qy      301  GTCCACGTGGGACGTGACGATCCCTTCAATGTGTCTCACTGTGGCCATGTATCCACC 360
Db      501  GTCCACGTGGGACGTGACGATCCCTTCAATGTGTCTCACTGTGGCCATGTATCCACC 560
Qy      361  GCCCTGTGAGCTCTGACCACTATAGAGGTGCAATGCGCGGACCTTCAATGGCCAGC 420
Db      561  GCCCTGTGAGCTCTGACCACTATAGAGGTGCAATGCGCGGACCTTCAATGGCCAGC 620
Qy      421  GTGTACAACAACGCGGACGCTGTGGGCTTGTGTGGGTGGCGCTGTCGACCAAGTTC 480
Db      621  GTGTACAACAACGCGGACGCTGTGGGCTTGTGTGGGTGGCGCTGTCGACCAAGTTC 680
Qy      481  TCCCTGCTGCTCTTCACTATCTGACAGCAATGTGTCCACCCGCGCTGAAGTGGCCAG 540
Db      681  TCCCTGCTGCTCTTCACTATCTGACAGCAATGTGTCCACCCGCGCTGAAGTGGCCAG 740
Qy      541  ATGCAGAACGACAGAGCTGCGGACGCGCACGCTGATGTATGCGGCTAAGCTGTGCGACA 600
Db      741  ATGCAGAACGACAGAGCTGCGGACGCGCACGCTGATGTATGCGGCTAAGCTGTGCGACA 800
Qy      601  CTGGCCACCTCTTACGCGCTGTGTGTCTCTCTCCGCGTCCGACGAGGACACGCCCCTG 660
Db      801  CTGGCCACCTCTTACGCGCTGTGTGTCTCTCTCCGCGTCCGACGAGGACACGCCCCTG 860
Qy      661  GACCGGAGACACGCGGCGCGCTGAGCGCCCTGGGACACAGGCTGTGTGGCCACCGTTGC 720
Db      861  GACCGGAGACACGCGGCGCGCTGAGCGCCCTGGGACACAGGCTGTGTGGCCACCGTTGC 920
Qy      721  AGCGAGTTTGGGCTTGTGACAGCCACCTATCTGATCTCTGCGGACACAGGCTCATCTC 780
Db      921  AGCGAGTTTGGGCTTGTGACAGCCACCTATCTGATCTCTGCGGACACAGGCTCATCTC 980
Qy      781  TCGCGAGGAGAGCGCGGTGAGACGACACTCTCTGGGCTCTACCTCTTGTGAAGATTTTC 840
Db      981  TCGCGAGGAGAGCGCGGTGAGACGACACTCTCTGGGCTCTACCTCTTGTGAAGATTTTC 1040
Qy      841  TCCAACTCTCTGCGCTTCTCCAGACGCTTGTGACACCACTTCTACCGCTACATGAGAC 900
Db      1041  TCCAACTCTCTGCGCTTCTCCAGACGCTTGTGACACCACTTCTACCGCTACATGAGAC 1100
Qy      901  CAGAGTTTCCCGACGACGCTCCAAAGGCTGATGAAAAAGCTGCGCGGAGCCGCGAC 960
Db      1101  CAGAGTTTCCCGACGACGCTCCAAAGGCTGATGAAAAAGCTGCGCGGAGCCGCGAC 1160
Qy      961  TGTCTCCCGGACCAATGCGGGGTGACAGAGGCTGGGGGTAG 1002
Db      1161  TGTCTCCCGGACCAATGCGGGGTGACAGAGGCTGGGGGTAG 1202

```

RESULT 4

AA\$42859
ID AA\$42859 standard; cDNA; 1778 BP.

AC AA\$42859;

DT 18-DEC-2001 (first entry)

DE Human G Protein-Coupled Receptor (GPCR) cDNA #54.

XX Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;
 KW attention deficit disorder; anxiety; depression; bipolar disorder; as;
 KW neurological disorder; Huntington's disease; dementia; obesity; anorexia;
 KW metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
 KW type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
 KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
 KW viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer;
 KW antidepressant; anorectic; PCR primer; gene therapy.

OS Homo sapiens.

XX WO200162797-A2.

```

XX      30-AUG-2001.
PD      23-FEB-2001; 2001WO-US005676.
PF      23-FEB-2001; 2001WO-US005676.
XX      23-FEB-2001; 2000US-0184247P.
PR      23-FEB-2001; 2000US-0184303P.
PR      23-FEB-2001; 2000US-0184304P.
PR      23-FEB-2001; 2000US-0184305P.
PR      23-FEB-2001; 2000US-0184397P.
PR      02-MAR-2001; 2000US-0186457P.
PR      03-MAR-2001; 2000US-0186810P.
PR      09-MAR-2001; 2000US-0188064P.
PR      13-MAR-2001; 2000US-0188880P.
PR      03-APR-2001; 2000US-0194344P.
PR      23-JUN-2001; 2000US-0213861P.
PR      11-JUL-2001; 2000US-0217369P.
PR      11-JUL-2001; 2000US-0217370P.
PR      14-JUL-2001; 2000US-0218337P.
PR      20-JUL-2001; 2000US-0218492P.
XX      (PHRA ) PHARMACIA & UPJOHN CO.
XX      Vogel I G, Wood LS, Parodi LA, Lind P,
PI      WPI: 2001-570628/64.
DR      P-PSDB; AAU25607.
XX      New isolated nucleic acid encoding a new G-protein coupled receptor
PT      polypeptide for detecting receptor modulators that can treat mental
PT      disorders, such as schizophrenia, anxiety, depression, or obesity.
XX      Claim 4; Page 90-91; 279pp; English.
XX      Sequences AA$42806-AA$42926 represent cDNA molecules and PCR primers for
XX      cDNA molecules encoding human G-protein coupled receptor (GPCR)
XX      polypeptides. The protein and DNA sequences of the invention can be used
XX      to identify compounds which bind to GPCR polypeptides and in screening
XX      for compounds that modulate GPCR activity. By screening a human subject
XX      for the presence of mutations in GPCR DNA, a GPCR-related disorder or a
XX      genetic predisposition can be diagnosed. The sequences can also be used
XX      for treatment and prevention of mental disorders such as schizophrenia,
XX      attention deficit disorder, anxiety, depression, dementia and bipolar
XX      disorder, neurological disorders such as Huntington's disease,
XX      Parkinson's disease and Tourette's syndrome, metabolic disorders such as
XX      obesity, anorexia and type 2 diabetes, cardiovascular disorders such as
XX      thrombosis, myocardial infarction, cardiomyopathy and atherosclerosis,
XX      viral infections caused by HIV and cancers
XX      Sequence 1778 BP; 376 A; 541 C; 498 G; 363 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1002; DB 4; Length 1778;
Best Local Similarity 100.0%; Pred.No. 3.8e-194;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1  ATGTGAGCTGACGCTGTTCAACGCGACAGGCTGTGTGAGAGACTGCTTCTGCCAG 60
Db      19  ATGTGAGCTGACGCTGTTCAACGCGACAGGCTGTGTGAGAGACTGCTTCTGCCAG 78
Qy      61  GACCTGACGCTGGGCTGTCACTGTTGTGCTGTGAGGCTGTGTGGTGGCGTGCACAG 120
Db      79  GACCTGACGCTGGGCTGTCACTGTTGTGCTGTGAGGCTGTGTGGTGGCGTGCACAG 138
Qy      121  GGCCTGTGTATCAACGCCCTGTGTGTGCGCAACTTACACAGACGAGCGCATGACC 180
Db      139  GGCCTGTGTATCAACGCCCTGTGTGTGCGCAACTTACACAGACGAGCATGACC 198
Qy      181  ATGCCGACGCTGACTTTTGTCAACATGCGAGTGGACGCTGTGTCTTACAGGCGCTGACC 240
Db      199  ATGCCGACGCTGACTTTTGTCAACATGCGAGTGGACGCTGTGTCTTACAGGCGCTGACC 258
Qy      241  CTTGTGACCTGTGCGGCCCCCGAGCTCCCGTGTGGGCTGTGAGTGTGGGCGGAA 300

```

Db 259 CTGTGACACTGCTGCGAGCCCCCGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAA 318
 QY 301 GTTCAGTGTGACATGACGATCCCTTTCAATGTGTCTTCATCTGTGGGCAATGTAATCTGACC 360
 Db 319 GTTCAGTGTGACATGACGATCCCTTTCAATGTGTCTTCATCTGTGGGCAATGTAATCTGACC 378
 QY 361 GGCCTGTGAGCTGTGACCACTACATGAGAGTGTGACATGTCGGGAGCACTACATGAGCAGC 420
 Db 379 GGCCTGTGAGCTGTGACCACTACATGAGAGTGTGACATGTCGGGAGCACTACATGAGCAGC 438
 QY 421 GTGTACAAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGCGCTGTAACAGCTTC 480
 Db 439 GTGTACAAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGCGCTGTAACAGCTTC 498
 QY 481 TCTCTGCTGTCTTTTCAATCTGTGAGCAATGTGTCCACCGCGCGCTAAGTGTGCGCAAG 540
 Db 499 TCTCTGCTGTCTTTTCAATCTGTGAGCAATGTGTCCACCGCGCGCTAAGTGTGCGCAAG 558
 QY 541 ATGCAAGACGACAGAGTGTGCGGACGCGGACGCTGTGTATTCGAGCTTACGTTGCGCAGCA 600
 Db 559 ATGCAAGACGACAGAGTGTGCGGACGCGGACGCTGTGTATTCGAGCTTACGTTGCGCAGCA 618
 QY 601 CTGGCCACCTCTACGCGCTGTGTACTCTCCCGCTCCGACGAGAGACACGCCCTTG 660
 Db 619 CTGGCCACCTCTACGCGCTGTGTACTCTCCCGCTCCGACGAGAGACACGCCCTTG 678
 QY 661 GACCGGACACGCGGCGGCTGTGAGCCCTCGGACACAGGCTGTGTGCGCACCGGTGTC 720
 Db 679 GACCGGACACGCGGCGGCTGTGAGCCCTCGGACACAGGCTGTGTGCGCACCGGTGTC 728
 QY 721 AGCAGTTTGGGCTGTGAGAGCGCACACTATCTGATCTGCGGGGACACAGTATCATC 780
 Db 739 AGCAGTTTGGGCTGTGAGAGCGCACACTATCTGATCTGCGGGGACACAGTATCATC 798
 QY 781 TCGCAGGAGGAGCCCGTGTGACGACACTACCTGGGCTACTGTGACTTTGAGAGATTTC 840
 Db 799 TCGCAGGAGGAGCCCGTGTGACGACACTACCTGGGCTACTGTGACTTTGAGAGATTTC 858
 QY 841 TCCAACTCTGTGCTTTTTCACAGAGCTTTGTGACACACTTTCTTACCGCTACATGAGAC 900
 Db 859 TCCAACTCTGTGCTTTTTCACAGAGCTTTGTGACACACTTTCTTACCGCTACATGAGAC 918
 QY 901 CAGAGCTTCCCGCAGCAAGCTCCAAAGGCTGATGAAAAAGCTGCGCGGGGACCGGAC 960
 Db 919 CAGAGCTTCCCGCAGCAAGCTCCAAAGGCTGATGAAAAAGCTGCGCGGGGACCGGAC 978
 QY 961 TGCTCCCGGACCACTGTGGGGTGTGACAGAGTGTGCGCTGAG 1002
 Db 979 TGCTCCCGGACCACTGTGGGGTGTGACAGAGTGTGCGCTGAG 1020

RESULT 5
 ABZ42598
 ID ABZ42598 standard; DNA; 1785 BP.
 XX
 AC ABZ42598;
 XX
 DT 04-MAR-2003 (first entry)
 XX
 DE Human G protein-coupled receptor MGC7035 nucleotide SEQ ID NO:679.
 DE
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KM G protein-coupled receptor; modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related disease; cell proliferation; autoimmunity; disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiovascular; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KM ulcer; gene; de.

OS Homo sapiens.
 XX
 PN MO200261087-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 19-DEC-2001, 2001WO-US050107.
 XX
 PR 19-DEC-2000, 2000US-0257144P.
 XX
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 PI Burner GC, Roush CL, Brown JP;
 DR WPI, 2003-046718/04.
 XX P-PSDB; ABP81753.
 XX
 PT New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX
 PS Disclosure; Fig 1; 523pp; English.
 XX
 CC The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorder, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 1785 BP; 381 A; 539 C; 501 G; 364 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1002; DB 8; Length 1785;
 Best Local Similarity 100.0%; Pred. No. 3.8e-194;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTGAGCTGCACTGTTTCAACGCGACAGGCTGTGTGAGAGAGTGTGCTGCTGCCAG 60
 Db 21 ATGTGAGCTGCACTGTTTCAACGCGACAGGCTGTGTGAGAGAGTGTGCTGCTGCCAG 80
 QY 61 GACCTGACCTGGGGGTGTCAGTGTGTGTGCTGGGCTGTGTGGGCGGCGCAGG 120
 Db 81 GACCTGACCTGGGGGTGTCAGTGTGTGTGCTGGGCTGTGTGGGCGGCGCAGG 140
 QY 121 GGCCTGTGTACAAACGCGCTGTGTGTGCTGCGCACTTACACAGCAAGGCGCAGATGACC 180
 Db 141 GGCCTGTGTACAAACGCGCTGTGTGTGCTGCGCACTTACACAGCAAGGCGCAGATGACC 200
 QY 181 ATGCGGACGCTGTACTTTTCAACATGAGAGTGTGCGAGCGCTGTGTCTACGCGCTGGCC 240
 Db 201 ATGCGGACGCTGTACTTTTCAACATGAGAGTGTGCGAGCGCTGTGTCTACGCGCTGGCC 260
 QY 241 CCGTGTGACACTGTGCGGCGCGCGAGCTCCGAGTGGGCGCTGTGTGAGTGTGGGCGGCGAA 300

```

Db      ||| CCTGTGACACCTGTGCGGCCCCCGAGCTCCCGTGGGGGCTGTGAGATGTGGGGGGA 320
Qy      ||| GTCAACGGGACATCGACAGATCCCTTCAATGTCTCTCACTGGTGGCCATGACTCCACC 360
Db      ||| GTCAACGGGACATCGACAGATCCCTTCAATGTCTCTCACTGGTGGCCATGACTCCACC 380
Qy      ||| GCCCTGTGAGCTCTGACCACTACATGAGAGTGTGCACTGGCGGGGACCTTCAATGGCCAGC 420
Db      ||| GCCCTGTGAGCTCTGACCACTACATGAGAGTGTGCACTGGCGGGGACCTTCAATGGCCAGC 440
Qy      ||| GTGTACAAACACGGGGACGTGTGGGCTTGTGTGGGGTGGCGGCTGTACCAAGCTTC 480
Db      ||| GTGTACAAACACGGGGACGTGTGGGCTTGTGTGGGGTGGCGGCTGTACCAAGCTTC 500
Qy      ||| TCCCTGCTGCTCTTCTCACTGTGAGGCTGTGTGCACTGGCGGGGACCTTCAATGGCCAGC 540
Db      ||| TCCCTGCTGCTCTTCTCACTGTGAGGCTGTGTGCACTGGCGGGGACCTTCAATGGCCAGC 560
Qy      ||| ATGAGAAACGACAGAGCTGCGGACGCGGCTGTGTGCACTGGCGGGGACCTTCAATGGCCAGC 600
Db      ||| ATGAGAAACGACAGAGCTGCGGACGCGGCTGTGTGCACTGGCGGGGACCTTCAATGGCCAGC 620
Qy      ||| CTGGCCACCTCTTACAGGCTGTGTGCTACTCTCCCGGCTCCGACGGAGACACGCGCCCTG 660
Db      ||| CTGGCCACCTCTTACAGGCTGTGTGCTACTCTCCCGGCTCCGACGGAGACACGCGCCCTG 680
Qy      ||| GACCGGGACACGGGGCGGCTGTGAGCCCTGGGACACAGAGGCTGTGTGCGGACCGCTGTG 720
Db      ||| GACCGGGACACGGGGCGGCTGTGAGCCCTGGGACACAGAGGCTGTGTGCGGACCGCTGTG 740
Qy      ||| AGCGAGTTGGGCTGTGAGCGGACACTATCTGATCTGTGGGGGACACGCTCATCATC 780
Db      ||| AGCGAGTTGGGCTGTGAGCGGACACTATCTGATCTGTGGGGGACACGCTCATCATC 800
Qy      ||| TCGCAGGGGAAAGCCGCTGTGACGACACTACCTGGGGCTACTGACCTTTGTGAAGATTTC 840
Db      ||| TCGCAGGGGAAAGCCGCTGTGACGACACTACCTGGGGCTACTGACCTTTGTGAAGATTTC 860
Qy      ||| TCCAACTCTGTGGCTTCTTCAGCAGCTTGTGACACCACTTCTCAACGCTACATGAAC 900
Db      ||| TCCAACTCTGTGGCTTCTTCAGCAGCTTGTGACACCACTTCTCAACGCTACATGAAC 920
Qy      ||| CAGAGCTTCCCAAGAGCTTCCAAACGCGCTGATGAATAAAGCTCCCTGCGGGACCGGAC 960
Db      ||| CAGAGCTTCCCAAGAGCTTCCAAACGCGCTGATGAATAAAGCTCCCTGCGGGACCGGAC 980
Qy      ||| TCTCTCCCGGACCACTGTGGGGTGTGACGAGGCTGTGGCGTAG 1002
Db      ||| TCTCTCCCGGACCACTGTGGGGTGTGACGAGGCTGTGGCGTAG 1022

RESULT 6
AD030022
ID      AD030022 standard; cDNA; 1785 BP.
AC      AD030022;
XX      29-JUL-2004 (first entry)
DE      Human GPCR PGR8 polynucleotide, SEQ ID NO:1124.
KW      G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW      transgenic mouse; neurological disorder; adrenal gland disorder;
KW      colon disorder; intestinal disorder; cardiovascular disorder;
KW      muscular disorder; blood disorder; immune disorder; bone disorder;
KW      joint disorder; metabolic disorder; nutritive disorder; cancer;
KW      kidney disorder; liver disorder; lung disorder; breast disorder;
KW      ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW      skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW      thymus disorder; thyroid disorder; antiparkinsonian; autonomic;
KW      cytostatic; antiinflammatory; vasodilator; antinutritional; antirhythmic;
KW      CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;

```

```

KW      virnucide; hepatotropic; antibacterial; antinaemic; antieborrheic;
KW      dermatological; antilucer; antihydroid; antiallergic; anorectic;
KW      immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW      gene; ss.
XX      Homo sapiens.
XX      MO2004040000-A2.
XX      13-MAY-2004.
XX      09-SEP-2003; 2003WO-US028226.
XX      09-SEP-2002; 2002US-0409303P.
XX      09-APR-2003; 2003US-0461329P.
XX      (PRIM-) PRIMAL INC.
XX      Galenaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,
XX      Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
XX      WPI; 2004-390329/36.
XX      P-PSDB; AD029606.
XX      Novel mammalian G protein coupled receptors, useful for identifying
XX      compounds that modulates diagnosing and treating disease condition
XX      associated with GPCR dysfunction e.g. autoimmune diseases, angina
XX      pectoris, Parkinson's disease.
XX      Claim 151; SEQ ID NO 1124; 542bp; English.
XX      The invention relates to human and mouse G protein-coupled receptors
XX      (GPCRs) and nucleic acids encoding them. The invention also relates to
XX      sequences at least 90% identical to the GPCR proteins and nucleic acids
XX      of the invention; methods of treating, preventing or diagnosing diseases
XX      associated with GPCRs of the invention; methods of screening for
XX      compounds useful in the treatment of GPCR-related diseases; a transgenic
XX      mouse comprising a GPCR gene of the invention; a mouse comprising a
XX      mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
XX      from the transgenic mice; kits comprising several mice, each of which has
XX      a mutation in a different GPCR gene of the invention; and kits comprising
XX      probes which hybridise to GPCR polynucleotides of the invention. The
XX      invention further discloses variants of the GPCR polypeptides and vectors
XX      comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
XX      be used in the diagnosis, treatment or prevention of a wide variety of
XX      diseases including neurological disorders (e.g., Alzheimer's disease,
XX      depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
XX      disorders of the adrenal gland; disorders of the colon or intestine
XX      (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
XX      syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
XX      myocardial infarction); muscular disorders; blood disorders (e.g.,
XX      anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
XX      AIDS); bone and joint disorders (e.g., osteoarthritis, Rheumatoid
XX      arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
XX      obesity, enzyme deficiency-related diseases or vitamin deficiency-related
XX      diseases); and disorders of the kidney, liver, breast, ovary,
XX      uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
XX      thyroid (e.g., cancer). The present sequence represents a GPCR-encoding
XX      nucleic acid of the invention. Note: The full sequence data for this
XX      patent did not form part of the printed specification; those sequences
XX      not shown were obtained in electronic format directly from WIPO at
XX      ftp.wipo.int/pub/published_poc_sequences.
XX      Sequence 1785 BP; 381 A; 539 C; 501 G; 364 T; 0 U; 0 Other;
XX      SO

Query Match      100.0%; Score 1002; DB 12; Length 1785;
Best Local Similarity 100.0%; Pred. No. 3.8e-194;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 ATGTGAGCTGTGACGCTGTTCACAGGACGCTGTGTGAGAGCTGCTGCTGCCAG 60
Db      21 ATGTGAGCTGTGACGCTGTTCACAGGACGCTGTGTGAGAGCTGCTGCTGCCAG 80

```

```

QY 61 GACCTGACGCTGGGCTGTCACCTGTTGTGCTGCTGGGCTGGTGGGCGTGGCCAGTG 120
DB 81 GACCTGACGCTGGGCTGTCACCTGTTGTGCTGCTGGGCTGGTGGGCGTGGCCAGTG 140
QY 121 GGCCTGTGCTACACGCGCTGTGCTGTGGCCAACTACACAGCAAGGCGCAGCATGACC 180
DB 141 GGCCTGTGCTACACGCGCTGTGCTGTGGCCAACTACACAGCAAGGCGCAGCATGACC 200
QY 181 ATGCCGAGAGCTGATCTTTGTGCAACATGCGAGTGGCAGAGGCTGTGCTCAGCGCTGGTGC 240
DB 201 ATGCCGAGAGCTGATCTTTGTGCAACATGCGAGTGGCAGAGGCTGTGCTCAGCGCTGGTGC 260
QY 241 CCTGTGACCTGTGCTGGGCGCGCGAGCTCCCGGTGGGCGGTGGAGTGTGGGCGCGGAA 300
DB 261 CCTGTGACCTGTGCTGGGCGCGCGAGCTCCCGGTGGGCGGTGGAGTGTGGGCGCGGAA 320
QY 301 GTTCAGTGGGCACTGAGATCCCTTCAATGTGTCTCACTGTGGTGCATGTACTCACC 360
DB 321 GTTCAGTGGGCACTGAGATCCCTTCAATGTGTCTCACTGTGGTGCATGTACTCACC 380
QY 361 GCGCTGTGAGCTGCACTACATGAGCGTGCACCTGCGGAGCTACATGAGCCAGC 420
DB 381 GCGCTGTGAGCTGCACTACATGAGCGTGCACCTGCGGAGCTACATGAGCCAGC 440
QY 421 GTGTAAACACGGGGGACCGTGGGCGCTTCTGTGGGGTGGCGCGCTGTGTAACACTTC 480
DB 441 GTGTAAACACGGGGGACCGTGGGCGCTTCTGTGGGGTGGCGCGCTGTGTAACACTTC 500
QY 481 TCCTGCTGCTCTTCTACATCTGTGACGCACTATGTGCAACCGCGCGCTAGATGCGCAG 540
DB 501 TCCTGCTGCTCTTCTACATCTGTGACGCACTATGTGCAACCGCGCGCTAGATGCGCAG 560
QY 541 ATGCAGAACGAGAGCTGCGGACGCGCACTGTGTTCTCATCGTGTGTCAGCA 600
DB 561 ATGCAGAACGAGAGCTGCGGACGCGCACTGTGTTCTCATCGTGTGTCAGCA 620
QY 601 CTGGGCAACCTCTACACCGCTGTGGTACTCTCCCGGTCCGAGGAGGACAGCGCCCTG 660
DB 621 CTGGGCAACCTCTACACCGCTGTGGTACTCTCCCGGTCCGAGGAGGACAGCGCCCTG 680
QY 661 GACCGGAGACAGGCGCGGCTGTGAGCGCTGTGAGCAACAGGCTGTGGTGCACCGTGTG 720
DB 681 GACCGGAGACAGGCGCGGCTGTGAGCGCTGTGAGCAACAGGCTGTGGTGCACCGTGTG 740
QY 721 AGCGAGTTTGGGCTGTGAGCGCACTATCTGATCTCTGTGGGAGCAACGCTCATC 780
DB 741 AGCGAGTTTGGGCTGTGAGCGCACTATCTGATCTCTGTGGGAGCAACGCTCATC 800
QY 781 TGCAGAGGAGAGCGCGCTGTGAGCGCACTATCTGATCTCTGTGGGAGCAACGCTCATC 840
DB 801 TGCAGAGGAGAGCGCGCTGTGAGCGCACTATCTGATCTCTGTGGGAGCAACGCTCATC 860
QY 841 TCCAACTCTGTGCTCTTCTCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAA 900
DB 861 TCCAACTCTGTGCTCTTCTCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAA 920
QY 901 CAGAGCTTCCCGAGCAAGCTCCAAAGGCTGTGATGAAAAAGCTGCTGCGGAGACCGGAC 960
DB 921 CAGAGCTTCCCGAGCAAGCTCCAAAGGCTGTGATGAAAAAGCTGCTGCGGAGACCGGAC 980
QY 961 TGTCTCCCGAGCAACATGAGGAGTGTGACAGAGTGTGAGCTGAG 1002
DB 981 TGTCTCCCGAGCAACATGAGGAGTGTGACAGAGTGTGAGCTGAG 1022

```

RESULT 7
ACCT0841 Standard; mRNA; 1926 BP.

ID ACC70841;
XX ACC70841;
AC ACC70841;
XX
DT 20-NOV-2003 (first entry)
XX

```

DE Human G-protein coupled receptor 901 mRNA.
XX Human; anorectic; antidiabetic; antilipemic; hypothalamus;
KW G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
KM clobophobia; anorexia nervosa; gene; 88.
XX Homo sapiens.
FH Key
FT CDS Location/Qualifiers
FT 154..1155
FT /tag=a
FT /product="G-protein coupled receptor 901"
XX MO200303936-A1.
XX 17-APR-2003.
XX 02-OCT-2002; 2002WO-JP010250.
XX 02-OCT-2001; 2001JP-00306872.
XX (SUMU) SUMITOMO PHARM CO LTD.
XX Suguru E, Tsuchida A, Yamanaka M, Taiji M;
XX WPI, 2003-354886/33.
XX P-PSDB; ABR56303.
XX Inhibitors of expression or activity of G-protein coupled receptor 901
XX for treatment of lifestyle-related diseases and clobophobia.
PS Claim 2; Page 62-64; 91pp; Japanese.
XX The present invention relates to novel remedies for the treatment of
XX diseases containing an active component an inhibitor of the expression
XX or activity of hypothalamus-expressed G-protein coupled receptor 901 and
XX for treatment of clobophobia containing an active component a
XX potentiator of the expression or activity of G-protein coupled receptor
XX 901. The diseases which can be treated include obesity, diabetes and
XX hyperlipaemia, and clobophobia (anorexia nervosa). The present human
XX sequence is the mRNA for G-protein coupled receptor 901
XX
SQ Sequence 1926 BP; 407 A; 590 C; 546 G; 0 T; 383 U; 0 Other;
Query Match 100.0%; Score 1002; DB 8; Length 1926;
Best Local Similarity 80.7%; Pred. No. 3.8e-194;
Matches 809; Conservative 193; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGCTGACGCTGTTCAACGGCAACAGGCTGTGGAGAGGCTGCTGCTGGCCAG 60
DB 154 ATGTGAGCTGACGCTGTTCAACGGCAACAGGCTGTGGAGAGGCTGCTGCTGGCCAG 213
QY 61 GACCTGACGCTGGGCTGTCACCTGTTGTGCTGCTGGGCTGGTGGGCGTGGCCAGTG 120
DB 214 GACCTGACGCTGGGCTGTCACCTGTTGTGCTGCTGGGCTGGTGGGCGTGGCCAGTG 273
QY 121 GGCCTGTGCTACACGCGCTGTGCTGTGGCCAACTACACAGCAAGGCGCAGCATGACC 180
DB 274 GGCCTGTGCTACACGCGCTGTGCTGTGGCCAACTACACAGCAAGGCGCAGCATGACC 333
QY 181 ATGCCGAGAGCTGATCTTTGTGCAACATGCGAGTGGCAGAGGCTGTGCTCAGCGCTGGTGC 240
DB 334 ATGCCGAGAGCTGATCTTTGTGCAACATGCGAGTGGCAGAGGCTGTGCTCAGCGCTGGTGC 393
QY 241 CCTGTGACCTGTGCTGGGCGCGCGAGCTCCCGGTGGGCGGTGGAGTGTGGGCGCGGAA 300
DB 394 CCTGTGACCTGTGCTGGGCGCGCGAGCTCCCGGTGGGCGGTGGAGTGTGGGCGCGGAA 453
QY 301 GTTCAGTGGGCACTGAGATCCCTTCAATGTGTCTCACTGTGGTGCATGTACTCACC 360
DB 454 GTTCAGTGGGCACTGAGATCCCTTCAATGTGTCTCACTGTGGTGCATGTACTCACC 513
QY 361 GCGCTGTGAGCTGCACTACATGAGCGTGCACCTGCGGAGCTACATGAGCCAGC 420

```



```

Db      514 GCCUCUCUAGCCUCCACCAUCAUACGAGCGUGGACCGCGGAGCCUACUAGCCAGC 573
QY      421 GGTATACAAACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGCGCTGTGACAGCTTC 480
Db      574 GGGUACAAACCGCGACCGUGCGGCUUUGUGGGUGGGCGGCUUACCAAGUUC 633
QY      481 TCTCGCTGCTCTTCAATCGACATGCGACATGTGTCCACCCCGCGCTAGAGTGGCCAG 540
Db      634 UCCUGGCUUCUUAUACUUGACCAUUGGUCACCCCGCGCUAGAUUGCGCCAG 693
QY      541 ATGCGAAACGAGAACTGCCGACGCGCTGTGTTCATCGGCTAGTGTGTGACGACA 600
Db      694 AVGCGAAACGAGAACTGCCGACGCGCTGTGTTCATCGGCTAGTGTGTGACGACA 753
QY      601 CTGGCCACCTCTTACGCGCTGTGTGTCTCTCCGCGCTCGGACGAGAGACAGCCCTG 660
Db      754 CUGGCACTCCUUAUACGCGGUGGUCUACUCCCGCGUCGCGAGGAGACAGCCCTG 813
QY      661 GACCGGAGACAGCGGCGCGCTGAGCGCTCGGACACAGGCTGTGTGTGTGTGTGTC 720
Db      814 GACCGGAGACAGCGGCGCGCTGAGCGCTCGGACACAGGCTGTGTGTGTGTGTGTC 873
QY      721 ACGAGTTTGGGCTGTGACCGCACTATCTATCTGCTGTGGGACACAGCTATCATC 780
Db      874 ACGAGUUGGGGCUUAGACCGCACACUACUAGUCCUGGGGACACAGUACUAC 933
QY      781 TCGGAGGAGAGCGCGCTGAGACGACACACTATCTATCTGCTGTGGGACACAGCTATC 840
Db      934 UCGGAGGAGAGCGCGCTGAGACGACACACTATCTATCTGCTGTGGGACACAGCTATC 993
QY      841 TCCAACTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db      994 UCCAAACUCCUGGCGUUCUACGAGAGUUGUACACCAUUCUACCCUACCAUAGAC 1053
QY      901 CAGAGTTTCCCGCAGCAAGCTCCAAACGCTGTATGAAAAAGTCCCTGCGGAGACCGGAC 960
Db      1054 CAGAGUUCUCCCGCAGCAAGCTCCAAACGCTGTATGAAAAAGTCCCTGCGGAGACCGGAC 1113
QY      961 TGCTCCCCGAGACCAATGAGGGGTGACAGAGTCTGCTGCTGAG 1002
Db      1114 UGCUCGCCGAGCACCAUAGGGGUGGACAGAGUUCUUGG 1155

```

RESULT 8

ACa61176
ID ACa61176 standard; cDNA, 1928 BP.

AC ACa61176;

DT 11-AUG-2003 (first entry)

DE cDNA encoding human G-protein coupled receptor 23992.

Human; ss; gene; G-protein coupled receptor; gene therapy; pneumonia;
immune disorder; Crohn's disease; Grave's disease; respiratory disorder;
diphtheria; haematopoietic disorder; leukaemia; systemic sclerosis;
prostatic disorder; benign prostate hypertrophy; tumour; breast disease;
acute mastitis; Paget's disease; muscular disorder; rhabdomyosarcoma;
neurological disorder; cerebral oedema; Parkinson's disease; atrophy;
blood vessel disorder; atherosclerosis; testicular disease; syphilis;
epididymal disease; thyroid disease; hyperthyroidism; cretinism; AIDS;
kidney disorder; cystic renal dysplasia; glomerulonephritis;
cardiovascular disease; heart failure; pericarditis; pancreatic disease;
pancreatitis; diabetes mellitus; thymus disease; thymic hypoplasia;
Hodgkin disease; spleen disease; splenomegaly; Gaucher disease;
liver disease; hepatic failure; alcoholic liver disease; T-cell disorder;
systemic lupus erythematosus; B-cell disorder; Burkitt Lymphoma;
multiple myeloma; platelet disorder; thrombocytopenia;
haemolytic-uraemic syndrome.

OS Homo sapiens.

XX

```

FH      Key      Location/Qualifiers
FT      CDS      160..1161
FT      FT      /tag= a
FT      FT      /product= "GPCR 23992"
FT      FT      /note= "G-protein coupled receptor"
XX      US2003017539-A1.
XX      PD      23-JAN-2003.
XX      PF      07-JUN-2002; 2002US-00165844.
XX      PR      02-JUN-1998; 98US-00088857.
XX      PR      02-SEP-1998; 98US-00145745.
XX      PR      21-JAN-1999; 99US-00234923.
XX      PR      02-JUN-1999; 99US-00324465.
XX      PR      28-JUN-1999; 99US-00340880.
XX      PR      26-AUG-1999; 99US-00383745.
XX      PR      16-DEC-1999; 99US-00464685.
XX      PR      18-DEC-2000; 2000US-00741783.
XX      PA      (MILL-) MILLENNIUM PHARM INC.
XX      PI      Glucksmann MA, Hodge MR, Hunter JU, Rudolph-Owen LA;
XX      PI      Slios-Santiago I, Welch NS;
XX      DR      WPI; 2003-401672/38.
XX      DR      P-Psdb; AB099900.
XX      PT      New nucleic acid molecule encoding a G-protein coupled receptor
XX      PT      polypeptide, e.g., 2871, 14926, 17723 or 23992, useful for diagnosing
XX      PT      and/or treating cancer or immune, respiratory, hematologic or
XX      PT      cardiovascular disorders.
XX      PS      Claim 2; Fig 21; 149pp; English.
XX      CC      The invention relates to an isolated G-protein coupled receptor nucleic
XX      CC      acid molecule encoding the G-protein coupled receptors 2871, 14926, 17723
XX      CC      and 23992. The nucleic acid molecule is useful in monitoring, diagnosing
XX      CC      and treating immune disorders (e.g. Crohn's disease and Grave's disease);
XX      CC      respiratory disorders (e.g. pneumonia or diphtheria); haematopoietic
XX      CC      disorders (e.g. leukaemia and systemic sclerosis); prostatic disorders
XX      CC      (e.g. benign prostate hypertrophy and tumours); breast diseases (e.g.
XX      CC      acute mastitis and Paget's disease); muscular disorders (e.g.
XX      CC      rhabdomyosarcoma); neurological disorders (e.g. cerebral oedema and
XX      CC      Parkinson's disease); disorders involving blood vessels (e.g.
XX      CC      atherosclerosis); diseases involving testis and epididymis (e.g.
XX      CC      kidney disorders (e.g. cystic renal dysplasia and glomerulonephritis);
XX      CC      cardiovascular diseases (e.g. heart failure and pericarditis); pancreatic
XX      CC      diseases (e.g. pancreatitis and diabetes mellitus); diseases involving
XX      CC      the thymus (e.g. thymic hypoplasia and Hodgkin disease); diseases
XX      CC      involving the liver (e.g. splenomegaly and Gaucher disease); diseases
XX      CC      -cell disorders (e.g. hepatic failure and alcoholic liver disease); T
XX      CC      cell disorder (e.g. Burkitt lymphoma and multiple myeloma); platelet disorders
XX      CC      (e.g. thrombocytopenia and haemolytic-uraemic syndrome). The nucleic
XX      CC      acids may also be used in chromosome mapping, tissue typing,
XX      CC      pharmacogenomics and forensic biology, and as surrogate markers. The
XX      CC      present sequence represents cDNA encoding the human G-protein coupled
XX      CC      receptor 23992
XX      SQ      Sequence 1928 BP; 408 A; 588 C; 547 G; 385 T; 0 U; 0 Other;
XX      Query Match      100.0%; Score 1002; DB 8; Length 1928;
XX      Best Local Similarity      100.0%; Pred. No. 3.ee-194;
XX      Matches 1002; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
QY      1 ATGTGAGACTGAGCTGTGTTCAACGGACAGGGCTGTGTGAGAGCTGCTGCTGCGAC 60
Db      160 ATGTGAGACTGAGCTGTGTTCAACGGACAGGGCTGTGTGAGAGCTGCTGCTGCGAC 219
QY      61 GACTGACGCTGGGGCTGTCACTGTGTGCTGTGGGCTGTGTGTGGGCTGTGTGACAGTG 120

```



```

Db      220 GACCTGACGCTGGGGCTGTACCTGTGTCTGCTGCTGGGCTGGTGGTGGGCGTGCACAGT 279
Qy      121 GGGCTGTGTACACACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      280 GGGCTGTGTGTACACACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 339
Qy      181 ATGCGGAGCTGTATCTTTGTCAACATGGGAGTGGGAGGCTGGTGGTCAAGCGCTTGGCC 240
Db      340 ATGCGGAGCTGTATCTTTGTCAACATGGGAGTGGGAGGCTGGTGGTCAAGCGCTTGGCC 399
Qy      241 CTTGTGCACTGTGTGGGCGCGCGAGCTCCCGGTGGGCGGTGTGTGTGTGTGTGTGTGTGT 300
Db      400 CTTGTGCACTGTGTGGGCGCGCGAGCTCCCGGTGGGCGGTGTGTGTGTGTGTGTGTGTGT 459
Qy      301 GTCCAGTGTGCACTGTGAGATCCCTTCAATGTGTCTCACTGTGTGGTGGTGTATCTCACC 360
Db      460 GTCCAGTGTGCACTGTGAGATCCCTTCAATGTGTCTCACTGTGTGGTGGTGTATCTCACC 519
Qy      361 GCGCTGTGTGAGCTTGTGACCACTATGTAGAGCTGTGCGGAGGAGCTTATGTGTGTGTGT 420
Db      520 GCGCTGTGTGAGCTTGTGACCACTATGTAGAGCTGTGCGGAGGAGCTTATGTGTGTGTGT 579
Qy      421 GTGTACACACAGGCGGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db      580 GTGTACACACAGGCGGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 639
Qy      481 TCCTGTGTGTCTTCTTCAATCTGTGAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db      640 TCCTGTGTGTCTTCTTCAATCTGTGAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 699
Qy      541 ATGCAAGAACGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db      700 ATGCAAGAACGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
Qy      601 CTGGGCAACCTCTTACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db      760 CTGGGCAACCTCTTACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 819
Qy      661 GACCGGAGACAGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db      820 GACCGGAGACAGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 879
Qy      721 AGCGAGTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db      880 ACCAGTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 939
Qy      781 TCGGAGGAGGAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db      940 TCGGAGGAGGAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 999
Qy      841 TCCAAACTCTGTGAGCTTCTTCAGAGAGTTTGTGACACACTTCTCTTACCGCTACATGAA 900
Db      1000 TCCAAACTCTGTGAGCTTCTTCAGAGAGTTTGTGACACACTTCTCTTACCGCTACATGAA 1059
Qy      901 CAAAGCTTCTTCCGCAAGAGCTTCCACAGGCTGTATGAAAAAGCTGCTGCGGAGACCGGAC 960
Db      1060 CAAAGCTTCTTCCGCAAGAGCTTCCACAGGCTGTATGAAAAAGCTGCTGCGGAGACCGGAC 1119
Qy      961 TGTCTCCCGGAGACATGTGGGAGTGTGACAGAGAGTGTGCGGTAG 1002
Db      1120 TGTCTCCCGGAGACATGTGGGAGTGTGACAGAGAGTGTGCGGTAG 1161

```

RESULT 9

AAZ90175
ID AAZ90175 standard; cDNA; 1614 BP.

XX AC AAZ90175;
XX AC
DT 19-MAY-2000 (first entry)
XX Human chemokine receptor nucleotide sequence.

```

XX Chemokine receptor; interleukin-8 compound inhibitor; chromosome 7p22;
KW inflammation; wound healing; neutropenia; myeloid leukaemia; tumour; 89;
KW toxin delivery; hypermegakaryocytopenic disease; polycythaemia vera.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 129..1127
FT /tag= a
FT /product= "Chemokine receptor"
FT /transl_except= (Pos:159..161, aa:Xaa)
FT /transl_except= (Pos:177..179, aa:Xaa)
FT /note= "Xaa = Unknown"
XX
XX MO200000515-A2.
XX
XX 06-JAN-2000.
XX
XX 29-JUN-1999; 99WO-US012829.
XX
XX 29-JUN-1998; 98US-00106800.
XX
XX 22-JAN-1999; 99US-00236166.
XX
XX (HYSE-) HYSEQ INC.
XX
XX WPI; 2000-170907/15.
XX
XX P-PSDB; AAY78856.
XX
XX New nucleic acid encoding chemokine receptor, useful for diagnosis and
XX treatment of e.g. neutropenia, inflammation and leukemia.
XX
XX Claim 1, Fig 1, 138pp; English.
XX
XX This sequence represents a human chemokine receptor coding sequence. The
XX sequence is derived from a sequence isolated from a human foetal liver-
XX spleen cDNA library. The chemokine receptor (see AAY78856) encoded by the
XX extended nucleotide sequence inhibits the activity of interleukin-8-type
XX compounds through competition for cell binding sites. The chemokine
XX receptor gene is located on the short arm of chromosome 7 at 7p22. The
XX hybridization probe or a PCR primer, the nucleotide sequence may also be
XX used for chromosome/gene mapping or in the recombinant production of
XX polypeptides and the production of antisense or triplex-forming molecules
XX for the control of gene expression. The chemokine receptor polypeptides
XX are used to raise specific antibodies, also for purification, detection
XX or modulation of interleukin-8-type chemokines (for diagnosis or
XX prognosis, or monitoring chemokine recruitment at a site of infection or
XX inflammation). The protein sequence can also be used as molecular weight
XX markers or food supplements, and to screen compound libraries for
XX specific binding agents, potential agonists or antagonists. Antibodies
XX raised against the chemokine receptor polypeptide sequence are used to
XX detect or purify the polypeptide, also for the diagnosis and treatment of
XX activated or inflamed cells or tissues, and to promote the healing of
XX wounds. The polypeptide and antibodies are also used to prevent
XX neutropenia (associated with chemotherapy or radiation treatment to
XX protect myeloid precursors), inflammation or other immune responses; also
XX conditions associated with hyperproliferation of progenitor cells (e.g.
XX some myelogenous leukaemias, polycythaemia vera and
XX hypermegakaryocytopenic diseases). The antibodies are potentially
XX useful therapeutically, e.g. to carry toxins to tumour cells
XX
XX Sequence 1614 BP; 307 A; 509 C; 468 G; 326 T; 0 U; 4 Other;
XX
XX Query Match 99.9%; Score 1000.8; DB 3; Length 1614;
XX Best Local Similarity 99.7%; Pred. No. 6.6e-194;
XX Matches 999; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTGAGCTGCAAGCTGTTCACACGACACAGGCTGTGTGAGAGAGTGTGCTGTCCAG 60
Db 129 ATGTGAGCTGCAAGCTGTTCACACGACACAGGCTGTGTGAGAGAGTGTGCTGTCCAR 188
Qy 61 GACCTGACGCTGGGCTGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120

```

```

Db      189  GACCTGAGCTGGGGCTGTCACTTTGTGCTGCTGAGCTTGTGTGG3TGG3CTTGGCAATG 248
Qy      121  GGCTTGTGTCAACAGCCCTGCTGTGTGGCAACTACAGAGAGAGCCAGATGACC 180
Db      249  GGCTTGTGTCAACAGCCCTGCTGTGTGGCAACTACAGAGAGAGCCAGATGACC 308
Qy      181  ATGCCGAGGTGTACTTTGTCAACATGGAGTGGAGAGCCCTGTGTCTCAAGCCCTGACC 240
Db      309  ATGCCGAGGTGTACTTTGTCAACATGGAGTGGAGAGCCCTGTGTCTCAAGCCCTGACC 368
Qy      241  CCTGTGACCTGTCTGGGCCCCCGAGCTCCCGTGGAGCTGTGGAGTGTGGAGGAGAA 300
Db      369  CCTGTGACCTGTCTGGGCCCCCGAGCTCCCGTGGAGCTGTGGAGTGTGGAGGAGAA 428
Qy      301  GTCCAGTGGGCACTGCAATCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCACC 360
Db      429  GTCCAGTGGGCACTGCAATCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCACC 488
Qy      361  GCCCTGTGAGGCTGTGACCACTACATCGAGGCTGCACTGGCGGAGACTCATGGCCAGC 420
Db      489  GCCCTGTGAGGCTGTGACCACTACATCGAGGCTGCACTGGCGGAGACTCATGGCCAGC 548
Qy      421  GTGTACACACAGCGGCAAGTGTGGGCTTGTGTGGGAGCGGCTGTGACCAAGCTTC 480
Db      549  GTGTACACACAGCGGCAAGTGTGGGCTTGTGTGGGAGCGGCTGTGACCAAGCTTC 608
Qy      481  TCCTGTGCTCTTCTTCAATCATGTGAGCCATGTGTCCACCCGCGGCTAAGATGCCCAAG 540
Db      609  TCCTGTGCTCTTCTTCAATCATGTGAGCCATGTGTCCACCCGCGGCTAAGATGCCCAAG 668
Qy      541  ATGAGAAACGAGAACTGCGGCAAGCGGCAAGCTGTGTTCATCGGCTCACTGATGCCAGCA 600
Db      669  ATGAGAAACGAGAACTGCGGCAAGCGGCAAGCTGTGTTCATCGGCTCACTGATGCCAGCA 728
Qy      601  CTGSCCAACCTCTTCTGAGGCTGTGTGTCTCTCCCGGCTCGGCAAGGAGAGACAGCCCTTG 660
Db      729  CTGSCCAACCTCTTCTGAGGCTGTGTGTCTCTCCCGGCTCGGCAAGGAGAGACAGCCCTTG 788
Qy      661  GACCGGGAGACAGGGGCGGCTGAGAGCCCTCGGCAACAGAGCTGTGTGGCAACCGGTGAC 720
Db      789  GACCGGGAGACAGGGGCGGCTGAGAGCCCTCGGCAACAGAGCTGTGTGGCAACCGGTGAC 848
Qy      721  ACCGAGTTTGGGCTGTGAGAGCCCACTATCTGATCTTCTGTGGGCAACCGGTCAATC 780
Db      849  ACCGAGTTTGGGCTGTGAGAGCCCACTATCTGATCTTCTGTGGGCAACCGGTCAATC 908
Qy      781  TCGCAGAGGAAAGCCGCTGAGAGCACTACTGAGGCTTACTGCACTTTGTGAAGATTTTC 840
Db      909  TCGCAGAGGAAAGCCGCTGAGAGCACTACTGAGGCTTACTGCACTTTGTGAAGATTTTC 968
Qy      841  TCCAACTCTGTGGCTTCTTCCAGAGCTTTGTGACCACTTCTTACCGCTACATGTAAC 900
Db      969  TCCAACTCTGTGGCTTCTTCCAGAGCTTTGTGACCACTTCTTACCGCTACATGTAAC 1028
Qy      901  CAGAGCTTCCCGAGCAAGCTTCCAGAGGCTGATGAAAAAGCTCCCTGCGGGAGCCGGGAC 960
Db      1029  CAGAGCTTCCCGAGCAAGCTTCCAGAGGCTGATGAAAAAGCTCCCTGCGGGAGCCGGGAC 1088
Qy      961  TGTCTCCCGGAGCAATGGGGGTGAGAGAGGTGTGGCGTATG 1002
Db      1089  TGTCTCCCGGAGCAATGGGGGTGAGAGAGGTGTGGCGTATG 1130

```

RESULT 10
AAS42836

ID AAS42836 standard; cDNA, 1750 BP.
XX

AC AAS42836;
XX

DT 18-DEC-2001 (first entry)
XX

DE Human G Protein-Coupled Receptor (GPCR) cDNA #31.

```

XX      XX      Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;
KW      KW      attention deficit disorder; anxiety; depression; bipolar disorder; ss;
KW      KW      neurological disorder; Huntington's disease; dementia; obesity; anorexia;
KW      KW      metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;
KW      KW      type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;
KW      KW      cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;
KW      KW      viral infection; immunosuppressant; neuroleptic; nootropic; tranquilizer;
KW      KW      antidepressant; anorectic; PCR primer; gene therapy.
OS      Homo sapiens.
XX      XX      MO200162797-A2.
PN      30-AUG-2001.
XX      XX      23-FEB-2001; 2001WO-US005676.
PF      23-FEB-2000; 2000US-0184247P.
PR      23-FEB-2000; 2000US-0184303P.
PR      23-FEB-2000; 2000US-0184304P.
PR      23-FEB-2000; 2000US-0184305P.
PR      23-FEB-2000; 2000US-0184397P.
PR      02-MAR-2000; 2000US-0186457P.
PR      03-MAR-2000; 2000US-0186810P.
PR      09-MAR-2000; 2000US-0188064P.
PR      13-MAR-2000; 2000US-0188880P.
PR      03-APR-2000; 2000US-0194344P.
PR      23-JUN-2000; 2000US-0213861P.
PR      11-JUL-2000; 2000US-0217369P.
PR      11-JUL-2000; 2000US-0217369P.
PR      14-JUL-2000; 2000US-0218337P.
PR      20-JUL-2000; 2000US-0218492P.
XX      XX      (PHAA ) PHARMACIA & UPJOHN CO.
PA      Vogeli G, Wood LS, Parodi LA, Lind P;
XX      XX      WPI; 2001-570628/64.
PI      P-PSDB; AAU25584.
DR      DR      New isolated nucleic acid encoding a new G-protein coupled receptor
PT      PT      polypeptide for detecting receptor modulators that can treat mental
PT      PT      disorders, such as schizophrenia, anxiety, depression, or obesity.
XX      XX      Claim 4; Page 81; 279pp; English.
XX      XX      Sequences AAS42806-AA542926 represent cDNA molecules and PCR primers for
XX      XX      cDNA molecules encoding human G-protein coupled receptor (GPCR)
XX      XX      polypeptides. The protein and DNA sequences of the invention can be used
XX      XX      to identify compounds which bind to GPCR polypeptides and in screening
XX      XX      for compounds that modulate GPCR activity. By screening a human subject
XX      XX      for the presence of mutations in GPCR DNA, a GPCR-related disorder or a
XX      XX      genetic predisposition can be diagnosed. The sequences can also be used
XX      XX      for treatment and prevention of mental disorders such as schizophrenia,
XX      XX      attention deficit disorder, anxiety, depression, dementia and bipolar
XX      XX      disorder, neurological disorders such as Huntington's disease,
XX      XX      Parkinson's disease and Tourette's syndrome, metabolic disorders such as
XX      XX      obesity, anorexia and type 2 diabetes, cardiovascular disorders such as
XX      XX      thrombosis, myocardial infarction, cardiomyopathy and atherosclerosis,
XX      XX      viral infections caused by HIV and cancers
XX      XX      Sequence 1750 BP; 336 A; 549 C; 516 G; 349 T; 0 U; 0 Other;
SQ

```

Query Match 99.8%; Score 1000.4; DB 4; Length 1750;
Best Local Similarity 99.9%; Pred. No. 8e-194;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1  ATGTGAGCTGACGCTGTGTTCAACGAGCAAGGGCTGTGAGAGAGCTGCTTGCCTGACAG 60
Db      139  ATGTGAGCTGACGCTGTGTTCAACGAGCAAGGGCTGTGAGAGAGCTGCTTGCCTGACAG 198
Qy      61  GACCTGACGCTGGGAGCTGTCACTGTTGTGCTGTGGGCTGTGGTGTGGGCGTGGCCAGTG 120

```

```

Db      199 GACCTGACGTGGGGGCTGTCACGTGTGTGCTGTGGGCGTGGGCGTGGCCAGT 258
Qy      121 GGGCTGTGCTACAAAGCCCTGCTGGTGTGGCCAACTTACACAGAGCCAGCATGAC 180
Db      259 GGGCTGTGCTACAAAGCCCTGCTGGTGTGGCCAACTTACACAGAGCCAGCATGAC 318
Qy      181 ATGCGGAGCGTGTACTTGTGCAATGAGGAGTGGGAGGCTGGGTCTCAGGCGCTGGCC 240
Db      319 ATGCGGAGCGTGTACTTGTGCAATGAGGAGTGGGAGGCTGGGTCTCAGGCGCTGGCC 378
Qy      241 CCTGTGACCTGTGCTGGGCCCCCGAGCTCCCGGTGGCGGTGTGAGTGTGGCGGAA 300
Db      379 CCGTGTGACCTGTGCTGGGCCCCCGAGCTCCCGGTGGCGGTGTGAGTGTGGCGGAA 438
Qy      301 GTCCAGTGGGACCTGCAAGATCCCTTCAATGTGTCTCACTGGTGGCATGTACTCAC 360
Db      439 GTCCAGTGGGACCTGCAAGATCCCTTCAATGTGTCTCACTGGTGGCATGTACTCAC 498
Qy      361 GGCCTGTGAGCCTCGGCACTACATGAGGAGTGCATGCGGAGCACTACATGAGGAG 420
Db      499 GGCCTGTGAGCCTCGGCACTACATGAGGAGTGCATGCGGAGCACTACATGAGGAG 558
Qy      421 GTGTACAAACAAGCGGCACTGTGCGGCTTGTGTGGGAGTGGCGGTGTGACCACTTC 480
Db      559 GTGTACAAACAAGCGGCACTGTGCGGCTTGTGTGGGAGTGGCGGTGTGACCACTTC 618
Qy      481 TCCTCGCTGCTCTTCTACATCTGCAAGCATGTGTCCACCGCGGTGAGTGGCCAG 540
Db      619 TCCTCGCTGCTCTTCTACATCTGCAAGCATGTGTCCACCGCGGTGAGTGGCCAG 678
Qy      541 ATGCGAAGCGAAGGTGCGGACGCGGCTGTGTCTACGCGTCTACGCGTGTGGCCAG 600
Db      679 ATGCGAAGCGAAGGTGCGGACGCGGCTGTGTCTACGCGTCTACGCGTGTGGCCAG 738
Qy      601 CTGGCCACCTCTACGCGGTGTGTGTCTCTCCGCGTCCGCGAGGAGGACAGCCCTTC 660
Db      739 CTGGCCACCTCTACGCGGTGTGTGTCTCTCCGCGTCCGCGAGGAGGACAGCCCTTC 798
Qy      661 GACCGGAGACAGCGGCGGCTGTGAGCCCTCGGCAACAAGGTGTGTGGCCACCTGTGTC 720
Db      799 GACCGGAGACAGCGGCGGCTGTGAGCCCTCGGCAACAAGGTGTGTGGCCACCTGTGTC 858
Qy      721 AGCGAATTTGGGCTGTGAGGCGCACTATCTGTGTGTGGGCAACAAGGTATCATTC 780
Db      859 AGCGAATTTGGGCTGTGAGGCGCACTATCTGTGTGTGGGCAACAAGGTATCATTC 918
Qy      781 TCSCGAGGAGAGCCGCTGAGGACGACACTACCTGGGCTACTGCACTTTGAGAGATTTC 840
Db      919 TCSCGAGGAGAGCCGCTGAGGACGACACTACCTGGGCTACTGCACTTTGAGAGATTTC 978
Qy      841 TCCAAACTCTGTGCTTCTTCAGACAGTTTGTGACACCACTTCTACCGCTACATGAC 900
Db      979 TCCAAACTCTGTGCTTCTTCAGACAGTTTGTGACACCACTTCTACCGCTACATGAC 1038
Qy      901 CAGAGCTTCCCGAGCAAGCTTCCAAAGGCTGTATGAAAAAGTGGCTTCGCGGAGCCGCGAC 960
Db      1039 CAGAGCTTCCCGAGCAAGCTTCCAAAGGCTGTATGAAAAAGTGGCTTCGCGGAGCCGCGAC 1098
Qy      961 TGTCTCCCGGACCAATGAGGAGTGCAGAGGCTGTGGCGTAG 1002
Db      1099 TGTCTCCCGGACCAATGAGGAGTGCAGAGGCTGTGGCGTAG 1140

```

```

RESULT 11
ABX73053
ID ABX73053 standard; cDNA; 1878 BP.
XX
AC ABX73053;
XX
XX 14-MAR-2003 (first entry)
XX
DE Human G-protein coupled receptor cDNA #3, Incyte clone 2214673CB1.

```

```

XX      XX Human; ss; gene: G-protein coupled receptor; GPCR; neoplastic disorder;
KW      KW neurological disorder; immune disorder; cytostatic; pancreatic cancer;
KW      KW follicular carcinoma of the thyroid; leiomyoma of the uterus; epilepsy;
KW      KW interstitial nephritis; cancer.
OS      OS Homo sapiens.
PN      PN US2002106655-A1.
XX      XX
PD      PD 08-AUG-2002.
XX      XX
PF      PF 28-JUN-2001; 2001US-00895686.
PR      PR 17-SEP-1998; 98US-00156513.
XX      XX
PA      PA (BAND/) BANDMAN O.
PA      PA (LALP/) LAL P G.
PA      PA (TANG/) TANG Y T.
PA      PA (BAUG/) BAUGHN M R.
XX      XX
PI      PI Bandman O, Lal PG, Tang YT, Baughn MR;
PT      PT
PT      PT WPI; 2002-697866/75.
DR      DR P-PDB; ABUS4867.
XX      XX
PT      PT New cDNAs encoding G protein coupled receptors are useful for the
PT      PT diagnosis, prognosis, treatment and evaluation of therapies for
PT      PT neoplastic, neurological and immune disorders.
XX      XX
PS      PS Claim 2; Page 30-31; 61pp; English.
XX      XX
CC      CC The invention relates to an isolated cDNA encoding G-protein coupled
CC      CC receptor (GPCR) appearing as ABUS4865-ABUS4870. Also included are
CC      CC fragments of the cDNAs, species variants having at least 75% identity to
CC      CC the cDNAs, vectors comprising the cDNAs, a host cells comprising the
CC      CC above vectors, producing a protein (comprising culturing the above host
CC      CC cell under expression conditions and recovering the protein), using a
CC      CC cDNA to detect expression of a nucleic acid in a sample or to screen for
CC      CC compounds or molecules which bind to the cDNAs, using the GPCR proteins
CC      CC to screen compounds or molecules for ligands, using a GPCR protein to
CC      CC prepare and purify antibodies, an anti-GPCR antibody and using the
CC      CC antibody to detect expression of a GPCR protein in a sample and its
CC      CC diagnosis of cancer. The invention is useful for the diagnosis,
CC      CC prognosis, treatment and evaluation of therapies for neoplastic,
CC      CC neurological and immune disorders, particularly follicular carcinoma of
CC      CC the thyroid, leiomyoma of the uterus, pancreatic cancer, epilepsy,
CC      CC interstitial nephritis and immune response as a complication of cancer.
CC      CC The present sequence encodes a human GPCR protein of the invention
XX      XX
SQ      SQ Sequence 1878 BP; 382 A; 585 C; 533 G; 378 T; 0 U; 0 Other;
XX      XX
Query Match 99.8%; Score 1000.4; DB 6; Length 1878;
Best Local Similarity 99.9%; Pred. No. 8.1e-194;
Matches 1001; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX      XX
Qy      Qy 1 ATGTGAGCTGCAAGCTGTTCAACGCGACAGGAGTGTGAGAGAGTGCCTGCTGCCAG 60
Db      139 ATGTGAGCTGCAAGCTGTTCAACGCGACAGGAGTGTGAGAGAGTGCCTGCTGCCAG 198
Qy      Qy 61 GACCTGCACTGGGCTGTGCACTGTTGTGCTGTGGGCTGTGGTGTGGGCGTGGCAG 120
Db      199 GACCTGCACTGGGCTGTGCACTGTTGTGCTGTGGGCTGTGGTGTGGGCGTGGCAG 258
Qy      Qy 121 GGGCTGTGCTACAAAGCCCTGCTGGTGTGGCCAACTTACACAGAGCCAGCATGAC 180
Db      259 GGGCTGTGCTACAAAGCCCTGCTGGTGTGGCCAACTTACACAGAGCCAGCATGAC 318
Qy      Qy 181 ATGCGGAGCGTGTACTTGTGCAATGAGGAGTGGGAGGCTGGGTCTCAGGCGCTGGCC 240
Db      319 ATGCGGAGCGTGTACTTGTGCAATGAGGAGTGGGAGGCTGGGTCTCAGGCGCTGGCC 378
Qy      Qy 241 CCGTGTGACCTGTGCTGGGCCCCCGAGCTCCCGGTGGCGGTGTGAGTGTGGCGGAA 300

```

Db	379	CTGTGACCTGCTCTGGGCCCCCGAGCTCCGGTGGGCGCTGTGAGTGTGGGGCGGAA	438
QY	301	GTCCAGTGGGCACTGCAGATCCCTTCAATGTGTCTCACTGGTGGGCATGTACTCCAC	366
Db	439	GTCACAGTGGCACTGCAGATCCCTTCAATGTGTCTCACTGGTGGGCATGTACTCCAC	499
QY	361	GCCCTGTGAGCCCTCCGACCACTACATGTAGAGCTGCATCTCCCGGACCTACATGGCCAG	420
Db	499	GCCCTGTGAGCCCTCCGACCACTACATGTAGAGCTGCATCTCCCGGACCTACATGGCCAG	558
QY	421	GTTGACCAACCGGGCACTGTGGGCTTGTGTGGGTGGGCGCGCTGCAGCACTTC	480
Db	559	GTTGACCAACCGGGCACTGTGGGCTTGTGTGGGTGGGCGCGCTGCAGCACTTC	618
QY	481	TCTCGCTGTCTTCTTCACTGTGAGCCATGTGTCCACCCGCGGTAGAGTCCGCAAG	540
Db	619	TCTCGCTGTCTTCTTCACTGTGAGCCATGTGTCCACCCGCGGTAGAGTCCGCAAG	678
QY	541	ATGCAGAACGCAAGAGTGCAGCGGCAAGCGGCGGTGTTCATCGGCTACGTGTGCCAGA	600
Db	679	ATGCAGAACGCAAGAGTGCAGCGGCAAGCGGCGGTGTTCATCGGCTACGTGTGCCAGA	738
QY	601	CTGGCCACCTCTTACGCGCTGTGTACTCTCCGGTCCGACAGGAGGACAAGCCCTG	660
Db	739	CTGGCCACCTCTTACGCGCTGTGTACTCTCCGGTCCGACAGGAGGACAAGCCCTG	798
QY	661	GACCGGGACAACGGGCGCGGCTGGAGCCCTGGGACACAGGCTCTGTGTGCCACGTTGTC	720
Db	799	GACCGGGACAACGGGCGCGGCTGGAGCCCTGGGACACAGGCTCTGTGTGCCACGTTGTC	858
QY	721	ACGCAAGTTTGGGCTCTGCAGCGCACACTATCTGATCCGTGGGGACACAGGTCAATATC	780
Db	859	ACGCAAGTTTGGGCTCTGCAGCGCACACTATCTGATCTGTGTGGGGACACAGGTCAATATC	918
QY	781	TGCGAAGGGAAGCCCTGTGACGCAACTACCTTGGGGCTACCTGCATTTGTGAAGATTC	840
Db	919	TGCGAAGGGAAGCCCTGTGACGCAACTACCTTGGGGCTACCTGCATTTGTGAAGATTC	978
QY	841	TCCAAACTCTGGGCTTCTCCGAGAGCTTTGTGACACACTTCTTACCGCTACATGAAC	900
Db	979	TCCAAACTCTGGGCTTCTCCGAGAGCTTTGTGACACACTTCTTACCGCTACATGAAC	1038
QY	901	CAGAGCTTCCCGAGCAAGCTTCAACGAGCTGTAAGAAAAGCTGGCCCTGGGGGACCGGAC	960
Db	1039	CAGAGCTTCCCGAGCAAGCTTCAACGAGCTGTAAGAAAAGCTGGCCCTGGGGGACCGGAC	1098
QY	961	TGCTCCCGAGCAATGGGGGTCCAGAGGTGCTGGGTAG	1002
Db	1099	TGCTCCCGAGCAATGGGGGTCCAGAGGTGCTGGGTAG	1140

```

QY 241 CTTGTGACCTGCTCGAGCCCCCGAGGCTCCCGGAGGCGCTGTGAGTGTGGGCGGAA 300
    |||
Db 378 CTTGTGACCTGCTCGAGCCCCCGAGGCTCCCGGAGGCGCTGTGAGTGTGGGCGGAA 437
QY 301 GTCCACGTGCACTGACAGATCCCTTCAATGTGTCTCACTGTGTGGCATGTATCCACC 360
    |||
Db 438 GTCCACGTGCACTGACAGATCCCTTCAATGTGTCTCACTGTGTGGCATGTATCCACC 497
QY 361 GCCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCGCGGACCTTACATATGGCAGC 420
    |||
Db 498 GCCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCGCGGACCTTACATATGGCAGC 557
QY 421 GTGTACAACACGCGGACGATGTGTGCGGCTGTGTGGGAGTGGCGGCTGTGACAGCTTC 480
    |||
Db 558 GTGTACAACACGCGGACGATGTGTGCGGCTGTGTGGGAGTGGCGGCTGTGACAGCTTC 617
QY 481 TCCCTGCTGCTTTCTTCAATCTGTGAGCCATGTGTCCACCGCGGCTATGAGTGGCCAG 540
    |||
Db 618 TCCCTGCTGCTTTCTTCAATCTGTGAGCCATGTGTCCACCGCGGCTATGAGTGGCCAG 677
QY 541 ATGCAGAACGCAAGAGTGTGCGGACGCAAGCTGTGTTCATCGGCTTACGTGTGTCCAGCA 600
    |||
Db 678 ATGCAGAACGCAAGAGTGTGCGGACGCAAGCTGTGTTCATCGGCTTACGTGTGTCCAGCA 737
QY 601 CTGGCCACCTCTTACGCGGCTGTGTCTACTCTCCCGGCTCGGACGAGGAGGACAGCCCTG 660
    |||
Db 738 CTGGCCACCTCTTACGCGGCTGTGTCTACTCTCCCGGCTCGGACGAGGAGGACAGCCCTG 797
QY 661 GACCGGAGCACGCGGCGGCTGTGAGCCCTTCGACACAGGCTGTGTGTGGCCAGCTGTGC 720
    |||
Db 798 GACCGGAGCACGCGGCGGCTGTGAGCCCTTCGACACAGGCTGTGTGTGGCCAGCTGTGC 857
QY 721 AGCGAGTTTGGGCTCTTGAAGCCACACTATCTGTCTCTGGGGGACACAGGTATATATC 780
    |||
Db 858 AGCGAGTTTGGGCTCTTGAAGCCACACTATCTGTCTCTGGGGGACACAGGTATATATC 917
QY 781 TCGGAGGAGGAGCCCGGAGGAGGACACTATCTGGGGGCTATGCACTTTGTAAGGATTC 840
    |||
Db 918 TCGGAGGAGGAGCCCGGAGGAGGACACTATCTGGGGGCTATGCACTTTGTAAGGATTC 977
QY 841 TCCAACTCTGTGCTTCTTCAAGAGCTTTGTGACACACTTCTTCACTCCGCTACATGAC 900
    |||
Db 978 TCCAACTCTGTGCTTCTTCAAGAGCTTTGTGACACACTTCTTCACTCCGCTACATGAC 1037
QY 901 CAGAAGCTTCCCGACGAAGCTTCGAAGGCTGATGAAAAAGCTGCGGAGACCGGAC 960
    |||
Db 1038 CAGAAGCTTCCCGACGAAGCTTCGAAGGCTGATGAAAAAGCTGCGGAGACCGGAC 1097
QY 961 TGTCTCCCGGACGACATGAGGAGGTGACAGAGGTGCTGCGCTAG 1002
    |||
Db 1098 TGTCTCCCGGACGACATGAGGAGGTGACAGAGGTGCTGCGCTAG 1139

RESULT 13
ACCT0860
ID ACCT0860 standard; DNA; 2216 BP.
XX
XX ACCT0860;
XX
XX 20-NOV-2003 (first entry)
XX
DE Insert cDNA sequence contained in pc901HISgalpha12.
XX
XX Human; anorectic; antidiabetic; antilipemic; hypothalamus;
XX G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
XX cibophobia; anorexia nervosa; ds.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
XX CDS 15..2102
XX FT /*tag= a
XX FT /product= "pc901HISgalpha12 protein"

```

```

XX
PN W02003030936-A1.
XX
XX 17-APR-2003.
XX
PF 02-OCT-2002; 2002MO-JP010250.
XX
PR 02-OCT-2001; 2001JP-00306872.
XX
XX (SUMU ) SUMITOMO PHARM CO LTD.
XX
PI Suguru E, Teuchida A, Yamanaka M, Taiji M;
XX
XX WPI: 2003-354866/33.
XX
XX P-PSDB; ABR56305.
XX
XX Inhibitors of expression or activity of G-protein coupled receptor 901
XX for treatment of lifestyle-related diseases and cibophobia.
XX
XX Example 5; Page 75-78; 91pp; Japanese.
XX
XX The present invention relates to novel remedies for the treatment of
XX diseases containing as an active component an inhibitor of the expression
XX or activity of hypothalamus-expressed G-protein coupled receptor 901 and
XX for treatment of cibophobia containing as an active component a
XX potentiators of the expression or activity of G-protein coupled receptor
XX 901. The diseases which can be treated include obesity, diabetes and
XX hyperlipaemia, and cibophobia (anorexia nervosa). The present sequence
XX was used to illustrate the invention
XX
SQ Sequence 2216 BP; 456 A; 688 C; 657 G; 415 T; 0 U; 0 Other;

```

```

Query Match 99.7%; Score 999; DB 8; Length 2216;
Best Local Similarity 100.0%; Pred. No. 1.6e-193;
Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGAGTGCAGAGCTGTTCACAGGACAGGCGTGTGAGAGAGCTGTGCTGCGAG 60
    |||
Db 15 ATGTGAGAGTGCAGAGCTGTTCACAGGACAGGCGTGTGAGAGAGCTGTGCTGCGAG 74
QY 61 GACCTGAGAGTGGGCTGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
    |||
Db 75 GACCTGAGAGTGGGCTGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 134
QY 121 GGCCTGTGTACACAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
    |||
Db 135 GGCCTGTGTACACAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 194
QY 181 ATGCCGAGCGTGTACTTTGTGACATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
    |||
Db 195 ATGCCGAGCGTGTACTTTGTGACATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 254
QY 241 CTTGTGACCTGCTCGAGCCCCCGAGGCTCCCGGAGGCGCTGTGAGTGTGGGCGGAA 300
    |||
Db 255 CTTGTGACCTGCTCGAGCCCCCGAGGCTCCCGGAGGCGCTGTGAGTGTGGGCGGAA 314
QY 301 GTCCACGTGCACTGACAGATCCCTTCAATGTGTCTCACTGTGTGGCATGTATCCACC 360
    |||
Db 315 GTCCACGTGCACTGACAGATCCCTTCAATGTGTCTCACTGTGTGGCATGTATCCACC 374
QY 361 GCCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCGCGGACCTTACATATGGCAGC 420
    |||
Db 375 GCCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCGCGGACCTTACATATGGCAGC 434
QY 421 GTGTACAACACGCGGACGATGTGTGCGGCTGTGTGGGAGTGGCGGCTGTGACAGCTTC 480
    |||
Db 435 GTGTACAACACGCGGACGATGTGTGCGGCTGTGTGGGAGTGGCGGCTGTGACAGCTTC 494
QY 481 TCCCTGCTGCTTTCTTCAATCTGTGAGCCATGTGTCCACCGCGGCTATGAGTGGCCAG 540
    |||
Db 495 TCCCTGCTGCTTTCTTCAATCTGTGAGCCATGTGTCCACCGCGGCTATGAGTGGCCAG 554
QY 541 ATGCAGAACGCAAGAGTGTGCGGACGCAAGCTGTGTTCATGAGGCTTACGTGTGTCCAGCA 600

```

```

Db      555 ATGAGAAACGAGAGCTGCGAGCGCAACGCTGCTTACGCTACGCTGCGCA 614
Qy      601 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGGCTCGGAGAGAGACGCCCC 660
Db      615 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGGCTCGGAGAGAGACGCCCC 674
Qy      661 GACCGGAGACAGCGGCGGCTGAGCGGCTCGGACACAGGCTGCTGCGGCACTGTC 720
Db      675 GACCGGAGACAGCGGCGGCTGAGCGGCTCGGACACAGGCTGCTGCGGCACTGTC 734
Qy      721 ACCGAGTTGGGCTCTGAGAGCGCACTATCTGATCTGCTGCGGAGACAGGTCATC 780
Db      735 ACCGAGTTGGGCTCTGAGAGCGCACTATCTGATCTGCTGCGGAGACAGGTCATC 794
Qy      781 TCGCGAGGAGAGCGGCTGAGAGCGCACTACTGCTGCGGCTGCTGCGGAGATTC 840
Db      795 TCGCGAGGAGAGCGGCTGAGAGCGCACTACTGCTGCGGCTGCTGCGGAGATTC 854
Qy      841 TCGAACTCTGAGCTCTCTCTGAGAGCGCACTACTGCTGCGGCTGCTGCGGAGATTC 900
Db      855 TCGAACTCTGAGCTCTCTCTGAGAGCGCACTACTGCTGCGGCTGCTGCGGAGATTC 914
Qy      901 CAGAGCTTCCCGAGCAAGCTTCCAGCGCTGATGAAAAAGCTGCGCGGAGACCGGCA 960
Db      915 CAGAGCTTCCCGAGCAAGCTTCCAGCGCTGATGAAAAAGCTGCGCGGAGACCGGCA 974
Qy      961 TGGTCCCGGAGACCACTGCGGCTGAGAGCGGCTGCGGCTGCGGCTGCGGCTGCGG 999
Db      975 TGGTCCCGGAGACCACTGCGGCTGAGAGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1013

```

```

RESULT 14
ACCT0859
ID ACC70859 standard; DNA; 2258 BP.
XX
AC ACCT0859;

```

```

XX 20-NOV-2003 (first entry)

```

```

XX Inset cDNA sequence contained in pc901HISgalpha16.

```

```

XX Human; anorectic; antidiabetic; antilipemic; hypothalamus;
XX G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
XX clobopobia; anorexia nervosa; ds.

```

```

XX Unidentified.

```

```

XX Key Location/Qualifiers
XX CDS 15..2159
XX /tag=a
XX /product="pc901HISgalpha16 protein"

```

```

XX MO2003030936-A1.
XX 17-APR-2003.
XX
XX 02-OCT-2002; 2002WO-JP010250.
XX
XX 02-OCT-2001; 2001JP-00306872.
XX
XX (SUMU) SUMITOMO PHARM CO LTD.
XX
XX Suguru E, Tsuchida A, Yamanaka M, Taiji M,
XX
XX MPI; 2003-354886/33.
XX
XX P-PSDB; ABR56304.

```

```

XX Inhibitors of expression or activity of G-protein coupled receptor 901
XX for treatment of lifestyle-related diseases and clobopobia.
XX
XX Example 5; Page 69-73; 91pp; Japanese.

```

```

CC The present invention relates to novel remedies for the treatment of
CC diseases containing as an active component an inhibitor of the expression
CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
CC for treatment of clobopobia containing as an active component a
CC potentiator of the expression or activity of G-protein coupled receptor
CC 901. The diseases which can be treated include obesity, diabetes and
CC hyperlipaemia, and clobopobia (anorexia nervosa). The present sequence
CC was used to illustrate the invention

```

```

XX Sequence 2258 BP; 436 A; 728 C; 687 G; 407 T; 0 U; 0 Other;

```

```

XX Query Match 99.7%; Score 999; DB 8; Length 2258;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-193;

```

```

XX Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGTGAGCTGAGCTGTTCAACGCAACAGGCTGTGTGAGAGAGCTGCTGCTGCA 60
Db      15 ATGTGAGCTGAGCTGTTCAACGCAACAGGCTGTGTGAGAGAGCTGCTGCTGCA 74
Qy      61 GACCTGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db      75 GACCTGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 134
Qy      121 GGCCTGCTACAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db      135 GGCCTGCTACAAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 194
Qy      181 ATGCGGAGCTGTACTTTTGTCAACATGGCACTGGAGGCTGTGCTGCTGCTGCTG 240
Db      195 ATGCGGAGCTGTACTTTTGTCAACATGGCACTGGAGGCTGTGCTGCTGCTGCTG 254
Qy      241 CCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db      255 CCGTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314
Qy      301 GTTCACTGTGAGCTGCAAGTCCCTTCAATGTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db      315 GTTCACTGTGAGCTGCAAGTCCCTTCAATGTGCTGCTGCTGCTGCTGCTGCTGCT 374
Qy      361 GGCCTGCTGAGCTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 420
Db      375 GGCCTGCTGAGCTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 434
Qy      421 GTTCACTGAGCTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 480
Db      435 GTTCACTGAGCTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 494
Qy      481 TCTTGTGCTGCTTCTTCACTGCTGCACTGCTGCACTGCTGCACTGCTGCACTGCT 540
Db      495 TCTTGTGCTGCTTCTTCACTGCTGCACTGCTGCACTGCTGCACTGCTGCACTGCT 554
Qy      541 ATGCAAGAACGAGAAAGCTGCGAGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db      555 ATGCAAGAACGAGAAAGCTGCGAGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 614
Qy      601 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGGCTCGGAGAGAGACGCCCC 660
Db      615 CTGGCCACCTCTACGCGCTGCTGCTACTCTCCCGGCTCGGAGAGAGACGCCCC 674
Qy      661 GACCGGAGACAGCGGCGGCTGAGCGGCTCGGACACAGGCTGCTGCGGCACTGTC 720
Db      675 GACCGGAGACAGCGGCGGCTGAGCGGCTCGGACACAGGCTGCTGCGGCACTGTC 734
Qy      721 ACCGAGTTGGGCTCTGAGAGCGCACTATCTGATCTGCTGCGGAGACAGGTCATC 780
Db      735 ACCGAGTTGGGCTCTGAGAGCGCACTATCTGATCTGCTGCGGAGACAGGTCATC 794
Qy      781 TCGCGAGGAGAGCGGCTGAGAGCGCACTACTGCTGCGGCTGCTGCGGAGATTC 840
Db      795 TCGCGAGGAGAGCGGCTGAGAGCGCACTACTGCTGCGGCTGCTGCGGAGATTC 854
Qy      841 TCGAACTCTGAGCTCTCTCTGAGAGCGCACTACTGCTGCGGCTGCTGCGGAGATTC 900

```

Db 855 TCCAACTCCTGGCCTTCTCCAGAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 914

QY 901 CAGAGCTTCCCGCAGCAAGCTTCCAAAGCTGATGAAAAAGTCCCTGCGGGGACCGGCAC 960

Db 915 CAGAGCTTCCCGCAGCAAGCTTCCAAAGCTGATGAAAAAGTCCCTGCGGGGACCGGCAC 974

QY 961 TGTCTCCCGGACCAATGAGGGGTGACAGAGTGTGGCG 999

Db 975 TGTCTCCCGGACCAATGAGGGGTGACAGAGTGTGGCG 1013

RESULT 15

ACC70861

ID ACC70861 standard; DNA; 2363 BP.

AC ACC70861;

DT 20-NOV-2003 (first entry)

XX Insert cDNA sequence contained in pc901HISgalphas2.

DE Human; anorectic; antidiabetic; antilipemic; hypothalamus;

KM G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;

KM cibophobia; anorexia nervosa; de.

XX Unidentified.

OS

XX Key Location/Qualifiers

FT CDS 15..2222

FT /tag= a

FT /product= "pc901HISgalphas2 protein"

XX WO2003030936-A1.

XX 17-APR-2003.

XX PD 02-OCT-2002; 2002WO-JP010250.

XX PF 02-OCT-2001; 2001JP-00306872.

XX PR (SUMU) SUMITOMO PHARM CO LTD.

XX PA

XX PI Sagaru E, Tsuchida A, Yamanaka M, Taiji M;

XX DR WPI; 2003-354886/33.

XX DR P-PSDB; ABR56306.

XX PT Inhibitors of expression or activity of G-protein coupled receptor 901

XX for treatment of lifestyle-related diseases and cibophobia.

XX PS Example 5; Page 81-84; 91pp; Japanese.

XX CC The present invention relates to novel remedies for the treatment of

CC diseases containing as an active component an inhibitor of the expression

CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and

CC for treatment of cibophobia containing as an active component a

CC potentiator of the expression or activity of G-protein coupled receptor

CC 901. The diseases which can be treated include obesity, diabetes and

CC hyperlipaemia, and cibophobia (anorexia nervosa). The present sequence

CC was used to illustrate the invention

XX

SQ Sequence 2363 BP; 522 A; 703 C; 674 G; 464 T; 0 U; 0 Other;

Query Match 99.7%; Score 999; DB 8; Length 2363;

Best Local Similarity 100.0%; Pred. No. 1.6e-193;

Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGCACTGTTCAACGACAGAGGCTGAGAGAGCTGCTGCTGCCAG 60

Db 15 ATGTGAGCTGCACTGTTCAACGACAGAGGCTGAGAGAGCTGCTGCTGCCAG 74

QY 61 GACCTGACAGTGGGGCTGTCATGTTGTGCTGCTGGCCTGTGAGGGCGTGCAGTG 120

Db 75 GACCTGACAGTGGGGCTGTCATGTTGTGCTGCTGGCCTGTGAGGGCGTGCAGTG 134

QY 121 GGCCTGTGCTACAAAGCCCTGCTGTGTGCTGGCCAACTTACACAGCAAGCCAGCATGACC 180

Db 135 GGCCTGTGCTACAAAGCCCTGCTGTGTGCTGGCCAACTTACACAGCAAGCCAGCATGACC 194

QY 181 ATGCCGGAAGTGTACTTGTGCAACATGGGAGTGGGAGGCGTGGTGTCAAGCCGCTGGCC 240

Db 195 ATGCCGGAAGTGTACTTGTGCAACATGGGAGTGGGAGGCGTGGTGTCAAGCCGCTGGCC 254

QY 241 CCTGTGCACTGTGCTGGCCGCCCGAGCTCCCGTGGGCGCTGTGAGTGTGGCGGCA 300

Db 255 CCTGTGCACTGTGCTGGCCGCCCGAGCTCCCGTGGGCGCTGTGAGTGTGGCGGCA 314

QY 301 GTCCACGTGGCACTGCAATGCCCTTCAATGTGTCTCACTGTGGCCATGTACTCCACC 360

Db 315 GTCCACGTGGCACTGCAATGCCCTTCAATGTGTCTCACTGTGGCCATGTACTCCACC 374

QY 361 GCCCTGTGAGCCTGACACCACTTACATGAGCGTGCATGCGCGGACCTTACATGGCCAGC 420

Db 375 GCCCTGTGAGCCTGACACCACTTACATGAGCGTGCATGCGCGGACCTTACATGGCCAGC 434

QY 421 GTGTACAAACAGCGGACAGTGTGCGGCTTGTGTGGGAGTGGCGCGCTGTGACCACTTTC 480

Db 435 GTGTACAAACAGCGGACAGTGTGCGGCTTGTGTGGGAGTGGCGCGCTGTGACCACTTTC 494

QY 481 TCTGTGCTCTCTTCTTACATCTGACAGCAATGTGTCCACCAGCGGCTGAGTGGCCAG 540

Db 495 TCTGTGCTCTCTTCTTACATCTGACAGCAATGTGTCCACCAGCGGCTGAGTGGCCAG 554

QY 541 ATGCAGAAAGCAAGAGTCCGACCGGACAGCTGTGTGTTATGATGGCTACCTGTGGCCAGCA 600

Db 555 ATGCAGAAAGCAAGAGTCCGACCGGACAGCTGTGTGTTATGATGGCTACCTGTGGCCAGCA 614

QY 601 CTGGCCACCTCTTACAGCGCTGTGTGCTACTCTCCCGGCTCGGAGGAGGACAGCCCTTG 660

Db 615 CTGGCCACCTCTTACAGCGCTGTGTGCTACTCTCCCGGCTCGGAGGAGGACAGCCCTTG 674

QY 661 GACCGGGACACGAGCGGCGGCTGTGAGCCCTGTGGGACACAGGCTGTGTGGCCACGTTGTC 720

Db 675 GACCGGGACACGAGCGGCGGCTGTGAGCCCTGTGGGACACAGGCTGTGTGGCCACGTTGTC 734

QY 721 AGCGAGTTGGGCTGTGAGCGGCACTATCTGATCTGTGGGAGGACAGGCTATCATTC 780

Db 735 AGCGAGTTGGGCTGTGAGCGGCACTATCTGATCTGTGGGAGGACAGGCTATCATTC 794

QY 781 TCCGAGGGAAGCCGCTGTGAGCGGCACTATCTGATCTGTGGGAGGACAGGCTATCATTC 840

Db 795 TCCGAGGGAAGCCGCTGTGAGCGGCACTATCTGATCTGTGGGAGGACAGGCTATCATTC 854

QY 841 TCCAACTCCTGGCCTTCTCCAGAGCTTGTGACACCACTTCTTACCGCTACATGAAC 900

Db 855 TCCAACTCCTGGCCTTCTCCAGAGCTTGTGACACCACTTCTTACCGCTACATGAAC 914

QY 901 CAGAGCTTCCCGCAGCAAGCTTCCAAAGCTGATGAAAAAGTCCCTGCGGGGACCGGCAC 960

Db 915 CAGAGCTTCCCGCAGCAAGCTTCCAAAGCTGATGAAAAAGTCCCTGCGGGGACCGGCAC 974

QY 961 TGTCTCCCGGACCAATGAGGGGTGACAGAGTGTGGCG 999

Db 975 TGTCTCCCGGACCAATGAGGGGTGACAGAGTGTGGCG 1013

Search completed: December 14, 2004, 22:43:51

Job time : 473.273 secs

This Page Blank (uspto)


```

Db      369  |||||CTGTGACCTGTGCGCCCCCGAGCTCCGGTGAGGCGGTGAGGTGTGGCGCGGAA 428
QY      301  |||||GTCCAGTGGCATGTGAGATCCCTTCAATGTCTCTCACTGTGGCCATGATCTCAACC 360
Db      429  |||||GTTCAGTGGCATGTGAGATCCCTTCAATGTCTCTCACTGTGGCCATGATCTCAACC 488
QY      361  |||||GCCCTGTGAGCTTCAGCACTATCATGAGCGTGACTGCGCGGAGCACTAATGGCCAGC 420
Db      489  |||||GCCCTGTGAGCTTCAGCACTATCATGAGCGTGACTGCGCGGAGCACTAATGGCCAGC 548
QY      421  |||||GTGTACAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGCGCTGTGAACAGCTTC 480
Db      549  |||||GTGTACAACACGCGGACGCTGTGCGGCTTGTGTGGGGTGGCGCGCTGTGAACAGCTTC 608
QY      481  |||||TCTGTGCTGTCTTCAATCTGCAACCATGTGTCAACCCGCGGCTAAGTGGCCAG 540
Db      609  |||||TCTGTGCTGTCTTCAATCTGCAACCATGTGTCAACCCGCGGCTAAGTGGCCAG 668
QY      541  |||||ATGCAGAACGAGAACTGCGGACGCGGACGCGTGTGTTATCGGCTAAGTGGCCAGCA 600
Db      669  |||||ATGCAGAACGAGAACTGCGGACGCGGACGCGTGTGTTATCGGCTAAGTGGCCAGCA 728
QY      601  |||||CTGGCCACCTCTTACGCGCTGTGTCTACTCTCCGCGTCCGAGGAGGACAACGCCCTTG 660
Db      729  |||||CTGGCCACCTCTTACGCGCTGTGTCTACTCTCCGCGTCCGAGGAGGACAACGCCCTTG 788
QY      661  |||||GACCGGGACACGCGGCGCGGCTGAGACCCCTGGGACACAGGCTGTGTGGCCACCGTGTGC 720
Db      789  |||||GACCGGGACACGCGGCGCGGCTGAGACCCCTGGGACACAGGCTGTGTGGCCACCGTGTGC 848
QY      721  |||||ACGCAATTTGGGCTGTGACGCGCACACTATGTACTCTGTGGGACACGCGTCAATC 780
Db      849  |||||ACGCAATTTGGGCTGTGACGCGCACACTATGTACTCTGTGGGACACGCGTCAATC 908
QY      781  |||||TGGCGAGGAAAGCCGTGAGACGCACTAATCTGGGGCTAATCGCACTTTGTGAAGATTTC 840
Db      909  |||||TGGCGAGGAAAGCCGTGAGACGCACTAATCTGGGGCTAATCGCACTTTGTGAAGATTTC 968
QY      841  |||||TCCAAACTCTGGGCTTCTCCAGAGCTTGTGACACCACTTCTACCGCTAATGAGAC 900
Db      969  |||||TCCAAACTCTGGGCTTCTCCAGAGCTTGTGACACCACTTCTACCGCTAATGAGAC 1028
QY      901  |||||CAGAGCTTCCCGAGCAAGCTCCCAAGCGCTGATGAAAAAGCTCCCTGCGGGAGCGGCAC 960
Db      1029  |||||CAGAGCTTCCCGAGCAAGCTCCCAAGCGCTGATGAAAAAGCTCCCTGCGGGAGCGGCAC 1088
QY      961  |||||TGGTCCCGGAGCAATGGGGGTGACGACGGTGTGGCGTAG 1002
Db      1089  |||||TGGTCCCGGAGCAATGGGGGTGACGACGGTGTGGCGTAG 1130

RESULT 2
US-09-409-778-1
; Sequence 1, Application US/09409778
; Patent No. 6472173
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Yeung, George
; TITLE OF INVENTION: A NOVEL CHEMOKINE RECEPTOR OBTAINED FROM
; FILE REFERENCE: A CDNA LIBRARY OF FETAL LIVER-SPLEEN
; CURRENT FILING DATE: 1999-09-22
; PRIOR FILING DATE: 1999-06-29
; PRIOR FILING DATE: 1999-06-29
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 09/106,800
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1

```

```

; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-409-778-1
Query Match      22.9%; Score 229.4; DB 4; Length 300;
Best Local Similarity 99.6%; Pred. No. 9,3e-43;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      772  |||||GTATCATCTGCGGAGGAAAGCCGCTGAGACGACACTATCTGGGGCTACTGCACTTTGTG 831
Db      1  |||||GTATCATCTGCGGAGGAAAGCCGCTGAGACGACACTATCTGGGGCTACTGCACTTTGTG 60
QY      832  |||||AAGATTTCTCCAAATCTGGGCTTCTCCAGAGCTTGTGACACCACTTCTCAACGC 891
Db      61  |||||AAGATTTCTCCAAATCTGGGCTTCTCCAGAGCTTGTGACACCACTTCTCAACGC 120
QY      892  |||||TACATGAACGAGCTTCCCGAGCAAGCTCCCAAGCGCTGATGAAAAAGCTGCCGCG 951
Db      121  |||||TACATGAACGAGCTTCCCGAGCAAGCTCCCAAGCGCTGATGAAAAAGCTGCCGCG 180
QY      952  |||||GACCGGCACTGCTCCCGGACCAATGGGGGTGACAGGTGTGGCGTAG 1002
Db      181  |||||GACCGGCACTGCTCCCGGACCAATGGGGGTGACAGGTGTGGCGTAG 231

RESULT 3
US-09-170-496D-3
; Sequence 3, Application US/09170496D
; Patent No. 655339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 655339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-3
Query Match      7.0%; Score 69.8; DB 4; Length 1089;
Best Local Similarity 42.4%; Pred. No. 8.9e-07;
Matches 380; Conservative 0; Mismatches 517; Indels 0; Gaps 0;

QY      52  |||||GCTGGCCAGACCTGACAGCTGGGGCTGTGACTGTTGTGCTGTGGGCTGTGTGGGC 111
Db      22  |||||GCTGGCCAGCTGGAATGCGCGGTGACCACTCTTTCCGCATCCCTTACATCTTTGTC 81
QY      112  |||||GTGCAAGTGGGCTGTGTGTAACAGCCCTGTGTGTGCGCAACTAACAAGAGCC 171
Db      82  |||||ATGCGCGTGGGCTGTGCAACCACTGCTGTGTGTGCGGCTTACCGCAAGTCAA 141
QY      172  |||||AGCATGACCATGCGGACGCTGACTTGTGCAACATGACAGTGGAGCGCTGTGAGC 231
Db      142  |||||CAGGCAACGAGCTGGGCGTGTACTGATGAACCTGACATGCGGACCTGTGATATC 201
QY      232  |||||GCCCTGGCCCTGTGTGACATGCTGTGGCCCCGAGGCTCCGGTGGGCGCTGTGAGTGTG 291
Db      202  |||||TGACAGCTGCGCTGTGGTGGTGAATCTTCTGTGACCAACAATGATCCAGGCC 261
QY      292  |||||GGCGGCAAGTCCAGTGGACCTGCAAGATCCCTTCAATGTGTCTCACTGTGGCCATG 351
Db      262  |||||GGGTCTGCAAGCTTTTGGGTTATCTTCAACCAATATCTACATCAGATGCGCTTC 321
QY      352  |||||TACTCCAGCGCTGTGAGCTTGACCACTAATGAGGTGACATGCGCGGAGCTTAC 411

```

Db 322 CTGTGCTGATCTCGGTGACCGGCTACTGTGGCCACCCACTCCGCTTGCCCGC 381
 Oy 412 ATGGCCAGGTGATCAACACCGGCACTGTGGGCTTGTGGGGTGGCGGCTGCTG 471
 Db 382 CTGGCGCGGTCAAGACCGCGCTGGCGGTAGCTCCGTGTGGGCAACGAGCTGGGC 441
 Oy 472 ACCAGCTTCTCTGCTGCTCTTCTACATCTGACGACATGTGTCAACCGCGGCTAGAG 531
 Db 442 GCCAATCGGCGCGCTGTTCCATGACGAGCTTTCGAGACCGCTCAACACACCTTC 501
 Oy 532 TGGCCAAAGATGCAAGACGAGAGCTGCCAGCCACGCTGTGTTTCATCGCTACGTC 591
 Db 502 TGTCTTGAAGAAAGTTCCCAATGAAAGCTGGGTGGCTGTGATGAACCTTATCGGGTTC 561
 Oy 592 GTGGCAGCACTGGCCACCTCTACCGGCTGTGCTACTCTCCGCGTCCGACGAGAGAC 651
 Db 562 GTGGGCTTCTTCTCCGTGGGCGCTCAATGCTGTCTGTGACCGGGCACTTGGCGGC 621
 Oy 652 ACCGCTGTGACCGGAGACGAGGCGGCTGGAGCCCTCGGACACAGGCTGTGTGGCC 711
 Db 622 GTGGCGGGGAGCGGTGTCCACCGGACCGCAGAGAGGCAAGATCAAGCGGCTGGCCCTC 681
 Oy 712 ACCGTGTGACGCGAGTTTGGGCTTGTGACGCGCACTATCTGATCTGTGTGGGCAACG 771
 Db 682 AGCTCATCGCCATCGTGTGTGTGTCTTGTGCGCCCTATCACTGTCTTGTGTCCGCG 741
 Oy 772 GTTCATCTCTCGGAGGAGGAGCGGCTGTGACGACACTACTGGGGCTACTGCACTTGTG 831
 Db 742 AGGCGCATCTACCTGTGGCGCGCCCTGTGGACTGTGGGCTTCAAGAGAGCGCGCTTTC 801
 Oy 832 AAGGATTTCTCCAACTCTGTGGCTTCTTCCAGAGAGCTTGTGACACACTTCTTCAACG 891
 Db 802 TACCAAGCTCACTGCTTGTACCAAGCTTCACTGTGTGTGGCGGACCCCACTCTTACTGC 861
 Oy 892 TACATGACCAAGAGCTTCCGACCAAGCTTCAACGCGCTGATGAAAAAGCTGCCCTGC 948
 Db 862 CTGGTCAACGAGGGCGCGCGACGATGTGGCAAGGCGCTGCAACACTGTCTCCG 918

RESULT 4
 US-09-016-434-1429
 ; Sequence 1429, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US

; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1429:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1365 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: 9598152
 ; US-09-016-434-1429

Query Match 7.0%; Score 69.8; DB 4; Length 1365;
 Best Local Similarity 42.4%; Pred. No. 9,4e-07;
 Matches 300; Conservative 0; Mismatches 517; Indels 0; Gaps 0;

Oy 52 GCCTGCAAGAGACTGTGACCTGGGCTGTGCACTGTGTGCTGCTGCTGCTGCTGCTGCTG 111
 Db 238 GGTGTCAGATGTGACCTGCGCGGTGACACACTCTTTCGCGCATCCCTTACATCTTGTG 297
 Oy 112 GTGCAAGTGGGCTGTGTCAACAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 171
 Db 298 ATGCGGCTGGGCTGTGCAACCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 357
 Oy 172 AGATATACATGCGGAGCGT 231
 Db 358 CAGCGCAAGAGTGTGGCGGTCTACCTGATGAACTTGAAGATGCGGACCTGTGTATATC 417
 Oy 232 GCGCTGAGCCCTGTGTGACCTGTGCGCCCGCCGAGCTCCCGGAGGCGCTGTGGAATGT 291
 Db 418 TGAAGCTGCGCTGT 477
 Oy 292 GGGGCGAAGTTCACGTGTGACATGACAGATCCCTTCAATGTGTCTTCACTGTGTGCGC 351
 Db 478 GGGTCTGCAAGGTCTTGTGGGTTCATCTTCAACCAATATATCAAGCATGGCTTTC 537
 Oy 352 TACTCCAGCGCTGT 411
 Db 538 CTGTGTGATCTCGGTGTGACCGGTACTGTGCTGTGCGCCACCACTTCCGCTTTCGCGC 597
 Oy 412 ATGGCAGGTGATCAACACGCGGACGCTGTGCGGCTGTGTGTGTGTGTGTGTGTGTGT 471
 Db 598 CTGGCGCGGTCAAGACCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657
 Oy 472 ACCAGCTTCTCTGCTGT 531
 Db 658 GCCAATCGGCGCGCTGT 717
 Oy 532 TGGCCAAAGATGCAAGACGAGAGCTGCCAGCGCACTGTGTGTGTGTGTGTGTGTGTGT 591
 Db 718 TGTCTTGAAGAAAGTTCCCAATGAAAGCTGGGTGGCTGTGATGAACCTTATCGGGTTC 777
 Oy 592 GTGGCAGCACTGGCGACCCCTTACCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 651
 Db 778 GTGGGCTTCTTCTCCGTGGGCGCTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 837
 Oy 652 ACCGCTGTGACCGGAGACGAGGCGGCTGTGAGCCCTCGGACACAGGCTGTGTGTGTGT 711
 Db 838 GTGGGAGGAGGAGGTGTCCACCGAGCGCAGAGAAAGCAAGATCAAGGGGCTGTGTGTGT 897
 Oy 712 ACCGTGTGACGCGAGTTTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 771
 Db 898 AGCTCATCGCCATCGT 957
 Oy 772 GTTCATCTCGGAGAGGAGAGCGCGGTGACGACACTACTGTGGGCTGTGTGTGTGTGTGT 831
 Db 958 AGGCGCATCTACTGTGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
 Oy 832 AAGGATTTCTCCAACTCTGTGGCTTCTCCAGAGAGCTTGTGTGACACACTTCTTCAACG 891
 Db 1018 TACCAAGCTCACTGTGCTTTCACAGCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 1077

QY 892 TACATGAACGAGCTTCCCGACAGCTCCACGCGCTGATGAAAAAGCTGCTCC 948
Db 1078 CTGGTCAACGAGGGGCGCCCGCAGCATGTGGCCAAAGGCCCTGCACACTGCTCCGC 1134

RESULT 5
US-09-170-496D-165
Sequence 165, Application US/09170496D
Patent No. 6555339
GENERAL INFORMATION:
APPLICANT: Behan, Dominic P.
APPLICANT: Chalmers, Derek T.
APPLICANT: Liaw, Chen W.
TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
TITLE OF INVENTION: Receptors
FILE REFERENCE: AREN-0040
CURRENT APPLICATION NUMBER: US/09/170,496D
CURRENT FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 294
SOFTWARE: Patentin version 3.1
SEQ ID NO 165
LENGTH: 1089
TYPE: DNA
ORGANISM: Homo sapiens
US-09-170-496D-165

Query Match 6.8%; Score 68.2; DB 4; Length 1089;
Best Local Similarity 42.3%; Pred. No. 2e-06;

Matches 379; Conservative 0; Mismatches 518; Indels 0; Gaps 0;

QY 52 GCCTCCAGAGACTGCAGCTGGGGCTGTCACTGTTGCTGCTGGGCTGCTGGGCTG 111
Db 22 GCGTCGACAGTGAAGCTGCGCGTGAACAACCTTTCCGCATCCCTTACATCTTGTG 81

QY 112 GTGCCAGTGGCGCTGTGCTACAGCGCGCTGTGCTGCGCAACCTACACAGCAAGCC 171
Db 82 ATGGGCGTGGGGCTGCGCCACCACTGCTGCTGTGGCGGCTTACCGCCAGGTGCA 141

QY 172 AGCATGACATGCGCGAGCTGTACTTGTACATGAGCAGTGGCAGGCTGTGTACG 231
Db 142 CAGGCAACGAGCGCGGCTGTACTGTATGATACCTGAGTCCGACCTGTGTATC 201

QY 232 GCCCTGGCGCTGTGACCTGTGCGCCCGGAGCTCCCGGTGGGCGCTGTGAGTGT 291
Db 202 TGCAAGTGGCGCTGTGAGTGTACTTCTGTGACCAACAACCTGATCCAGCGCC 261

QY 292 GGGCGGAGTGCAGCTGAGCTGACATGCTCCCTTCAATGTGTCTCACTGTGTGCT 351
Db 262 GGGTCTGCAAGCTCTTGGGTGTATTTCTACACCAATCTACATCAGCATGCGCTTC 321

QY 352 TACTCAGCGCGCTGTGAGCTGTGACCACTACATGAGGTGATGCTGCCGAGCTTAC 411
Db 322 CTGTGTGCAATCTGTGTGAGCGGCTACTGCTGTGCGCCACCACTCCCTTTCGCGC 381

QY 412 ATGGCGAGCTGTACACAGCGCGCAGTGTGGCGCTGTGAGGTGTGGCGGCTGTG 471
Db 382 CTGGCGCGCTGCAAGCGCGCGGTGAGCTGTGCTGTGTGTGAGGCAAGGAGCTGGC 441

QY 472 ACCAGCTTCTCTGCTGTCTTCTTACATGTGACATGATGTGTCACCGCGGCTGAG 531
Db 442 GCGAAGTGGCGCGCTGTGTGATGAGAGCTTTCGAGACCGCTACACACACTTC 501

QY 532 TGGCCCAAGATGCAAGAGCAAGCTGCGACCGCACTGTGTGTTCATGCTAGTGT 591
Db 502 TGCTTTAGAGTGTCCCATGAGAGGTGTGTGCTGTGAGTGTGAGTGTGAGTGTTC 561

QY 592 GTGCAAGAGTGGCAAGCTGTGTGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 651
Db 562 GTGGGCTTCTCTTCTGT 621

QY 652 ACGCCCTGAGCCGAGACAGCGGCGGCTGTGAGCTTTCGACACAGGCTGTGTGTGT 711

Db 622 GTGGGGGAGAGCTGTCCACCGAGCGCAGAGAGAAAGCCAAAGATCGCGGCTGCGCTC 681
QY 712 ACCGTGAGCAGCAAGTTGGCTGTGACCGCCACACTATCTGATCTGTGGGCAACG 771
Db 682 AGCCATTCGCGCATGT 741

QY 772 GTCATCATCTGGGAGAGAGCGGTGTGAGCAGCACTACTGTGGGCTGTGACTTGTG 831
Db 742 AGCGCATCTACTGTGGCGCGGCTGTGAGCTGTGCGGCTTTCAGAGAGCGCGTCTTTCGCA 801

QY 832 AAGATTTCTCCAACTCTGCGCTTCTTCAGAGCTTGTGACACCACTTCTTACCGC 891
Db 802 TACCAAGCTACTAGGCTTTCACCAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 861

QY 892 TACATGAACGAGCTTCCCGACAGCTCCACGCGCTGATGAAAAAGCTGCTCCGC 948
Db 862 CTGTCAACGAGGGGCGCCGAGCATGTGGCCAAAGGCCCTGCACACTGCTCCGC 918

RESULT 6
US-09-299-843A-65
Sequence 65, Application US/09299843A

Patent No. 6107475

GENERAL INFORMATION:

APPLICANT: Godiska, Ronald

APPLICANT: Gray, Patrick W.

APPLICANT: Schwaikart, Vicki L.

TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

ADDRESSEE: Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,843A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/088,337

FILING DATE: 01-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/153,848

FILING DATE: 17-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/977,452

FILING DATE: 17-NOV-1992

ATTORNEY/AGENT INFORMATION:

NAME: Jili E. Uhl

REGISTRATION NUMBER: 43,213

REFERENCE/DOCKET NUMBER: 27866/32059B

TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

TELEX:

INFORMATION FOR SEQ ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 2085 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 177..1310

US-09-299-843A-65

Query Match 6.3%; Score 62.8; DB 3; Length 2085;
 Best Local Similarity 43.7%; Pred. No. 3.9e-05;
 Matches 374; Conservative 0; Mismatches 477; Indels 5; Gaps 2;

118 GTGGGCTGTGCTCAACGCGCTGTGCTGAGCACTACAGAGGCGCAGATG 177
 |||||
 387 GTGGGCTGTGCTGCGCAACGGGCTGTGATCTACGATCTATTTCAGAGGCTCAAG 446
 |||||
 178 ACCATGCGGAGCTGTATCTTTGTCAATGAGCAAGGCTGTGCTGAGGCGCTG 237
 |||||
 447 ACCATGACGATACCTACCTGCTCACTGCGCGTGGGAGACATCTTTCTCTATT 506
 |||||
 238 GCCCTGTGACCTGTGCGCGCGCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGC 297
 |||||
 507 CTTCCTCT---CTGGGCTTACAGCGAAGCGAAGCTCTGTGATCTTTGGCGCTACCTGTG 562
 |||||
 298 GAAGTCCA-CGTGGCACTGAGATCCCTTCAATGTGCTCACTGGAGGCAATGTATCTC 356
 |||||
 563 TAAAGGATCTTTGGCATCTATTAAGTTAGCTTTTCAAGCGGATGCTGCTCTATG 622
 |||||
 357 CACCGCCCTGTGAGCTGCACTACATGAGCGTGAAGCTGCGCGGACCTTACATGCG 416
 |||||
 623 CATCAGATTGACCGGCTACGATGCTGACAGCGGCTGTGCGCTCATGCGCACCGCGC 682
 |||||
 417 CAGGCTTACACACCGGCGACGTTGCGGCTTCTGTGGGTGGCGCGCTGACAG 476
 |||||
 683 CCGCGTGTCTCTCATACGAGAGCTGTCTGTGGGATCTGAGTGTGCGCTCTTCT 742
 |||||
 477 CTTCCTGCTGTGCTCTTCAATCTGAGCATGTGTCAACCGGCGGCGCTGAGTGGC 536
 |||||
 743 CTTCATCTCCGAGGCTGCTTACAGCGGCTCTCAAGAGCGCGGAGAGACACGCTGAG 802
 |||||
 537 CAAGATGACAGACGAGAGCTGCGGACGCGCTGTGCTTCAATGCGCTACGTGTGCC 596
 |||||
 803 ATGCTCATGCTGTAGTGTCCCAAGGAGGCGCTTGTATCACCATCAAGTGGCGCAGATGT 862
 |||||
 597 AGCATGCGCACCTCTTACGCGGCTGTGCTACTTCCCGGCTCGGAGGAGACACGCGC 656
 |||||
 863 TTTGGGTCTCTAGTGTCTTATGCTGTGCTATGAGTTTGTGCTACTCATATCATCCGTAC 922
 |||||
 657 CTGGAACGGGAGACGCGGCGGCGTGGAGCCCTGGGACACAGAGTGTGGTGGCAGCGT 716
 |||||
 923 CTGTCTCAGGACGCAACTTTGAGCGGAAACAGGCGCATCAAGGTATCATATGCGGTGT 982
 |||||
 717 GTGCACGCACTTGTGGGCTGTGAGCGGACATATCTGATCTGTGGGAGACAGCGTCTAT 776
 |||||
 983 GGTAGTCTTATAGTCTTCCAGCTGCTTCAATGGGCTGTCTGGCTCAGACGTTGGC 1042
 |||||
 777 CATCTGCGAGGAGAGCGCGTGAAGCACTACCTGGGCTACTGCACTTTGTGAAGA 836
 |||||
 1043 CAATTCAMATCAACCAATAGAGCTCGAAGACAGACAGCTCAACATTTGCTATGA 1102
 |||||
 837 TTTCTCCAATCTCGGCTTCTCCAGAGCTTTGTGACACACACTTCTTCAACGCTCAT 896
 |||||
 1103 CGTCACTTCAAGCTGCGCTCTCGCTGCTGTGCTCAACCTTTCTTGAATGCTTCTAT 1162
 |||||
 897 GAACAGAGCTTCCCGGAGAGCTCAACGCGCTGATGAAAAAGCTGCTGCGGAGCG 956
 |||||
 1163 CGGCGTCAAGTTTCCGAGGACCTTCAAGCTTCTCAAGGACTTGGGCTGTCTCAGCCA 1222
 |||||
 957 GCACTGCTCCCGGAC 972
 |||||
 1223 GGAACGCTCCGGCAC 1238
 |||||

RESULT 7
 US-09-088-337B-65
 ; Sequence 65; Application US/09088337B
 ; Patent No. 6348574
 ; GENERAL INFORMATION:
 ; APPLICANT: Godtfrä, Ronald

Gray, Patrick W.
 Schweikart, Vicki L.
 TITLE OF INVENTION: No. 6348574e1 Seven Transmembrane Receptors
 NUMBER OF SEQUENCES: 66
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/088,337B
 FILING DATE: 01-Jun-1998
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/153,848
 FILING DATE: 17-NOV-1993
 APPLICATION NUMBER: US 07/977,452
 FILING DATE: 17-NOV-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6348574and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 31794
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 65:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2085 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 177..1310
 SEQUENCE DESCRIPTION: SEQ ID NO: 65:
 US-09-088-337B-65
 Query Match 6.3%; Score 62.8; DB 3; Length 2085;
 Best Local Similarity 43.7%; Pred. No. 3.9e-05;
 Matches 374; Conservative 0; Mismatches 477; Indels 5; Gaps 2;

118 GTGGGCTGTGCTCAACGCGCTGTGCTGAGCACTACAGAGGCGCAGATG 177
 |||||
 387 GTGGGCTGTGCTGCGCAACGGGCTGTGATCTACGATCTATTTCAGAGGCTCAAG 446
 |||||
 178 ACCATGCGGAGCTGTATCTTTGTCAATGAGCAAGGCTGTGCTGAGGCGCTG 237
 |||||
 447 ACCATGACGATACCTACCTGCTCACTGCGCGTGGGAGACATCTTTCTCTATT 506
 |||||
 238 GCCCTGTGACCTGTGCGCGCGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGC 297
 |||||
 507 CTTCCTCT---CTGGGCTTACAGCGAAGCGAAGCTCTGTGATCTTTGGCGCTACCTGTG 562
 |||||
 298 GAAGTCCA-CGTGGCACTGAGATCCCTTCAATGTGCTCACTGGAGGCAATGTATCTC 356
 |||||
 563 TAAAGGATCTTTGGCATCTATTAAGTTAGCTTTTCAAGCGGATGCTGCTCTATG 622
 |||||
 357 CACCGCCCTGTGAGCTGCACTACATGAGCGTGAAGCTGCGCGGACCTTACATGCG 416
 |||||
 623 CATCAGATTGACCGGCTACGATGCTGACAGCGGCTGTGCGCTCATGCGCACCGCGC 682
 |||||
 417 CAGCGTGTACACACCGGCGACGTTGCGGCTTCTGTGGGTGGCGCGCTGACAG 476
 |||||

```

Db      683 CCGGCTCTTCTCATGACGAGCTGTCGTGGGATCTGATCTGACCTCTTCT
Qy      477 CTCTCTCGCTGCTCTTCTTACATCTGACAGCCATGTCACCCCGCGCTAGAGTGC
Db      743 CTCATCTCCGAGCTGCTTACAGCGGCTCCAGAAAGAACAGCGGAGACACCGTGA
Qy      537 CAAGATGACAGACGAGAGCTCCGACGCAACGCGTGTCTCATCGGTACGTGTGC
Db      803 ATGCTCACTGTGATGAGGCCAAGTGAAGCTTTGATCAACATCAAGTGGCCAGATGT
Qy      597 AGACATGGCCACCTCTTACCGCGTGTGTCTTCTCCCGCTCCGACAGGAGACAGCC
Db      863 TTTTGGGTTCTAGTGCCTATGCTGTGATGATTTCTGCTCACTTATGATCCGAC
Qy      657 CCTGACCGGAGACACGCGGCGGTGAGCCCTGCGACACAGGCTGTGTGCGTCCGT
Db      923 CTGTGCTCAGGACCGAATTGAGCGAAGAACAGGCAATCAAGGATCATTTGCGGTGT
Qy      717 GTGACAGCATTTGGGCTGTGACGCAACATATCTGATCTGCGTGGGACACGCTGAT
Db      983 GGTAGCTTTCATATGCTTTCAGCTGCTCCCTCAATAGGGTGTCTGTGCTAGAGTGC
Qy      777 CATCTCGAGGAGAGCCGCTGAGACACTACTGCGGCTACTGCACTTTGTGAAGA
Db      1043 CACTTCAACATCAACATAGACGCTGGAAACAGACAGCACTCAACATTTGCTATGA
Qy      837 TTTCTCAAACTCTGAGCTTCTCCAGAGCTTTGTGACACTTCTCAACGCTATCAT
Db      1103 CGTCACTTACAGCTGCGCTCGTCCGTGCTGTGCTCAACCTTTCTTGTATGCTTCA
Qy      897 GAACAGAGCTTCCCGACAGCTTCAACGCGTGTATGAAGTGTGCGGAGACG
Db      1163 CGGCTCAAGTTCGCGAGCACTTTCAGACTTTCAGAGACTTGTGGCTGCTCAGCCA
Qy      957 GCACTGCTCCCGGAC 972
Db      1223 GGAACGGCTCCGAC 1238

```

RESULT 8

US-08-789-354-1

```

; Sequence 1, Application US/08789354
; Patent No. 5851798
; GENERAL INFORMATION:
; APPLICANT: Shabon, Usman
; APPLICANT: Bergsma, Derk
; TITLE OF INVENTION: Cloning of Human GPR14 Re
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Pasteo for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,354
; FILING DATE: 27-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: P50610
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-789-354-1

```

```

Query Match      5.8%; Score 58.4; DB 2; Length 2126;
Best Local Similarity 48.7%; Pred. No. 0.0038;
Matches 294; Conservative 0; Mismatches 286; Indels 24; Gaps 4;

```

```

Qy      118 GTGGGCTGTGCTTACAGCGCCCTGCTGTGTGCTGCACTTACAGAGAGGAGCATG 177
Db      883 GTGGGCTGTGCTTACAGCGCCCTTACAGCGCTGTGTGCTTACAGCGCTGTGTGCTG
Qy      178 ACCATGCGGAGCTGTGCTTGTGCAATGTCAGTGTGAGGCTGTGTGCTGAGCGCTG
Db      943 GCTTCATGTATGCTTACGTGTGCACTGAGCTGTGCGCACTGTGTGCTGCTGAGC
Qy      238 GCTTGTGCACTGTGTGCGCCCGGAGCTCCGCTGTGAGGCTGTGTGAGTGTGAGGCG
Db      1003 ATCCCTTCATGTGTGCGCACTTACAGAGAGTGTGCACTTGTGAGGAGTGTG
Qy      298 GAATGCGAGTGTGCACTGTGCAATTCCTTCAATGTGTCTACTGTGTGCTGTATGCC
Db      1063 CGGTGTCTTGTGCTGTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
Qy      358 ACCGCTGTGAGCTGTGCACTTACATGAGCTGTGCACTGTGCGGAGCTGTGCACTG
Db      1114 CTGACCGTATGAGCAGAGCGCTTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Qy      416 -----CAAGCTGTGCACTGTGCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT
Db      1174 GCGCCCAAGAGCTACCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Qy      472 ACCAGCTTCTGTGCTGTGCTTCTTACATGTGAGCTGTGTGTGTGTGTGTGTGTGT
Db      1234 ACGTGTGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Qy      532 TGGGCAAGTGTGAGAGCGAGAGCTGTGAGCGGAGCTGTGTGTGTGTGTGTGTGT
Db      1294 CCGGCTGTGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
Qy      589 GTGTGTGAGCACTGTGCACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
Db      1354 GCGGCGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Qy      643 AGGAGAGACGCGCTGTGAGCGGAGCAAGGCGGCGGCTGTGAGCGCTGTGAGCGG
Db      1414 TCGCAGCGGCGCTTCTTCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
Qy      703 CTGG 706
Db      1474 CTGG 1477

```

RESULT 9

US-09-110-937-1

```

; Sequence 1, Application US/09110937A
; Patent No. 6005074
; GENERAL INFORMATION:
; APPLICANT: SHABON, USMAN
; APPLICANT: BERGSMAN, DERK
; TITLE OF INVENTION: CLONING OF HUMAN GPR14 RECEPTOR
; FILE REFERENCE: P50610-1
; CURRENT APPLICATION NUMBER: US/09/110,937A
; CURRENT FILING DATE: 1998-07-06
; NUMBER OF SEQ ID NOS: 2

```


159 ACACAGAGAGCCAGCATGACATGCGGAGAGTGTACTTGTCAACATGGCAGTGGCAGG 218
153 GGGGATCACTTGGGCTTCACTTACCGGCGGTACCTGATCGATGACGTAACGAACA 212
219 CTTGTGTCTCAGGCGCCCTGGGCTTGTGCACTGTCTGGGCCCCCGAGCTCCGGGTGGG 278
213 ACCGGCCGATCTGGGCAAGGTGCTGTCAATGTGTGGGCGGCTTACAGAGAGCGCT 272
279 GCTGTGAGTGTGGGCGGCAAGTCCAGTGTGCACTGTGCAATCCCTTCAATGTGTCTC 338
273 GGTGACAGAGCTTGGGCGACCGCGGACGCTGTGTGTGCGGATGCGCAGAGACCGGA 332
339 ACTGTGTGCGCATATCTCCACCGGCTTGTGAGCTGTGACCATATCATGAGCTGTGACT 398
333 ACTGGGCAAGAGGTGCTACCGCTTGTGACAGAGCGGATGCGGCAACCGGGAACCT 392
399 GCGGCGGACCTATACATGAGCGACGCTGTACACACGCGGACGCTGTGTGTGGG 458
393 GAAGCGGCTGAGTACGTCCGCGCGGAGGTGGTGAAGAACTGGCGACAGGCGGCT 452
459 TGGCGGCTGTGACAGCTTCTTCTGTGTCTTGTACATGTGACGATGTGTCTAC 518
453 CGGATGTCTCTGTGGCGCTGGGCGGATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 512
519 CCGCGGCTGTGAGTGTGGCGGCAAGTGTGCAAGAGCGGAGCTGCGGCGGCTGTGT 578
513 GTTGTGCTTGT 572
579 CATCGGCTATGAGTGTGCGACATGTGGCACCCTTACCGGCTGTGTGT 626
573 GTGCTTCTTGT 620

RESULT 13

US-09-252-991A-6054/c
Sequence 6054, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074.788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094.190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 6054
LENGTH: 984
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6054

Query Match 5.7%; Score 57.6; DB 4; Length 984;
Best Local Similarity 44.3%; Pred. No. 0.00048;
Matches 234; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

99 CTTGTGTGTGGGCTGTCAGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 158
910 CTTGTGTGTGGGCTGTCAGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 851
159 ACACAGAGAGCCAGCATGACATGCGGAGAGTGTACTTGTCAACATGGCAGTGGCAGG 218
850 GGGGATCACTTGGGCTTCACTTACCGGCGGTACCTGATCGATGACGTAACGAACA 791
219 CTTGTGTCTCAGGCGCCCTGGGCTTGTGCACTGTCTGGGCCCCCGAGCTCCGGGTGGG 278
790 ACCGGCCGATCTGGGCAAGGTGCTGTCAATGTGTGGGCGGCTTACAGAGAGCGCT 731
279 GCTGTGAGTGTGGGCGGCAAGTGTGCAAGTGTGCAATGCCCTTCAATGTGTCTC 338

730 GGTGACAGAGTGTGGGCGGACCGCGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 671
339 ACTGTGTGCGCATATCTCCACCGGCTTGTGAGCTGTGACCATATGAGCGTGTGACT 398
670 ACTGGGCAAGAGGTGCTACCGCTTGTGACAGAGCGGATGTGCGGCAACCGGCGGAACCT 611
399 GCGGCGGACCTATATGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 458
610 GAAGCGGCTGAGT 551
459 TGGCGGCTGTGACAGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 518
550 CGGATGTCTCTGTGGCGCTGGGCGGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 491
519 CCGCGGCTGTGAGTGTGGCGGCAAGTGTGCAAGAGCGGAGCTGCGGACGCGTGTGT 578
490 GTTGTGCTTGT 431
579 CATCGGCTATGAGTGTGCGACATGTGGCACCCTTACCGGCTGTGTGT 626
430 GTGCTTCTTGT 383

RESULT 14

US-08-540-650B-6
Sequence 6, Application US/08540650B
Patent No. 6393325
GENERAL INFORMATION:
APPLICANT: HINUMA, Shuji
APPLICANT: FUJII, Ryo
APPLICANT: FUKUSUMI, Shoji
APPLICANT: OHTAKI, Tetsuya
APPLICANT: HOSODA, Masaki
APPLICANT: OHGI, Kazuhiko
APPLICANT: ONDA, Haruo
TITLE OF INVENTION: GALANIN RECEPTOR PROTEIN, PRODUCTION AND USE THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WATER STREET
CITY: BOSTON
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/540.650B
FILING DATE: 11-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 7-134412
FILING DATE: 31-MAY-1995
APPLICATION NUMBER: 6-326610
FILING DATE: 28-DEC-1994
APPLICATION NUMBER: 6-247599
FILING DATE: 13-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, DAVID S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45901
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1047
TYPE: Nucleic acid
STRANDEDNESS: Double

TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
IDENTIFICATION METHOD: S
US-08-540-650B-6

Query Match 5.5%; Score 55.2; DB 3; Length 1047;
Best Local Similarity 45.4%; Pred. No. 0.0017;
Matches 198; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

```
QY 89 CGCTGCTGGGCGCTGTGTGTGGGCGCTGCGCAATGAGGCTGTGTGTACAAAGCCCTGCTGTGC 148
DB 107 CGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 166
QY 149 TGGCCAACTTACACAGCAAGGCGAGCATGACATGCGGAGGTACTTGTCAATGG 208
DB 167 CCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 226
QY 209 CAGTGGCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
DB 227 ACTGAGCATGCGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 286
QY 269 CCCGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 328
DB 287 ACGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 346
QY 329 ATGTGTCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 388
DB 347 CCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 406
QY 389 AGCGTGACATGCGGCGGACCTTACATGTGCGGAGGTGTACAAACGCGGACGTGTGTGTGT 448
DB 407 CCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 466
QY 449 TCGTGTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 508
DB 467 TGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 526
QY 509 ATGTGTCCACCGCGGC 524
DB 527 TCTTCCACCGCGGC 542
```

RESULT 15

US-09-016-434-1423
Sequence 1423, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSER: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-4166
TELEFAX: (650) 845-0555
INFORMATION FOR SEQ ID NO: 1423:
SEQUENCE CHARACTERISTICS:
LENGTH: 1053 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 9559047
US-09-016-434-1423

Query Match 5.5%; Score 55.2; DB 4; Length 1053;
Best Local Similarity 45.4%; Pred. No. 0.0017;
Matches 198; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

```
QY 89 CGCTGCTGGGCGCTGTGTGTGGGCGCTGCGCAATGAGGCTGTGTGTACAAAGCCCTGCTGTGC 148
DB 107 CGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 166
QY 149 TGGCCAACTTACACAGCAAGGCGAGCATGACATGCGGAGGTACTTGTCAATGG 208
DB 167 CCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 226
QY 209 CAGTGGCAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
DB 227 ACTGAGCATGCGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 286
QY 269 CCCGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 328
DB 287 ACGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 346
QY 329 ATGTGTCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 388
DB 347 CCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 406
QY 389 AGCGTGACATGCGGCGGACCTTACATGTGCGGAGGTGTACAAACGCGGACGTGTGTGTGT 448
DB 407 CCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 466
QY 449 TCGTGTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 508
DB 467 TGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 526
QY 509 ATGTGTCCACCGCGGC 524
DB 527 TCTTCCACCGCGGC 542
```

Search completed: December 15, 2004, 02:56:16
Job time: 87.2892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:25:49 ; Search time 511.891 Seconds
(without alignments)
10792.424 Million cell updates/sec

Title: US-09-995-225b-1

Perfect score: 1002

Sequence: 1 atcgtgcagcgcagcgcgtc.....tcgcagcagtcgctgcgctag 1002

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*

19: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*

20: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*

21: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1002	100.0	1002	9	US-09-995-225-1
2	1002	100.0	1002	10	US-09-990-940-3
3	1002	100.0	1002	10	US-09-995-225-1
4	1002	100.0	1402	15	US-10-292-798-1015
5	1002	100.0	1561	15	US-10-017-161-1201
6	1002	100.0	1614	15	US-10-101-148-2
7	1002	100.0	1778	10	US-09-791-932-54
8	1002	100.0	1785	15	US-10-223-567A-679
9	1002	100.0	1928	14	US-10-165-844-14
10	1000.4	99.8	1750	10	US-09-791-932-31
11	1000.4	99.8	1878	9	US-09-895-686-9
12	1000.4	99.8	1893	16	US-10-296-115-435

13	577	57.6	2651	9	US-09-764-847-1782	Sequence 1782, Ap
14	577	57.6	2651	14	US-10-092-154-1782	Sequence 1782, Ap
15	358.6	35.8	349	10	US-09-918-995-8117	Sequence 8117, Ap
16	350.6	35.0	461	10	US-09-818-995-30890	Sequence 30890, A
17	304	30.3	497	9	US-09-895-686-59	Sequence 59, Appl
18	275	27.4	275	9	US-09-895-686-32	Sequence 32, Appl
19	236.4	23.6	238	9	US-09-895-686-30	Sequence 30, Appl
20	196	19.6	196	15	US-10-101-148-1	Sequence 1, Appl
21	178	17.8	930	9	US-09-764-847-1781	Sequence 1781, Ap
22	178	17.8	930	14	US-10-092-154-1781	Sequence 1781, Ap
23	169	16.9	294	9	US-09-895-686-61	Sequence 61, Appl
24	166	16.6	259	9	US-09-895-686-31	Sequence 31, Appl
25	162.6	16.2	259	9	US-09-895-686-60	Sequence 60, Appl
26	134	13.4	255	9	US-09-895-686-27	Sequence 27, Appl
27	122	12.2	281	9	US-09-895-686-29	Sequence 29, Appl
28	89.4	8.9	363	9	US-09-895-686-28	Sequence 28, Appl
29	69.8	7.0	1089	15	US-10-251-385-3	Sequence 3, Appl1
30	69.8	7.0	1089	15	US-10-267-811-1	Sequence 1, Appl1
31	69.8	7.0	1365	15	US-10-101-510-17	Sequence 17, Appl
32	69.8	7.0	1365	16	US-10-305-720-1429	Sequence 1429, Ap
33	69.8	7.0	1365	17	US-10-283-975A-113	Sequence 113, App
34	69.8	7.0	2693	10	US-09-850-948-1	Sequence 1, Appl1
35	69.8	7.0	2693	15	US-10-273-575-1	Sequence 1, Appl1
36	69.8	7.0	2696	15	US-10-225-567A-272	Sequence 272, App
37	69.8	7.0	2980	15	US-10-101-510-456	Sequence 456, App
38	68.2	6.8	1089	15	US-10-251-385-165	Sequence 165, App
39	62.8	6.3	1137	15	US-10-004-113-21	Sequence 21, Appl
40	62.8	6.3	2072	9	US-09-815-937-19	Sequence 19, Appl
41	62.8	6.3	2072	15	US-10-004-113-20	Sequence 20, Appl
42	60.8	6.1	1068	14	US-10-237-563-48	Sequence 48, Appl
43	58.4	5.8	1170	15	US-10-225-567A-537	Sequence 537, App
44	58.4	5.8	1170	17	US-10-775-920-14	Sequence 14, Appl
45	58.4	5.8	1215	15	US-10-220-883-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1

US-09-995-225-1

Sequence 1, Application US/09995225

Publication No. US20020193584A1

GENERAL INFORMATION:

APPLICANT: Chen, Ruoping

APPLICANT: Chu, Zhi Liang

APPLICANT: Dang, Huang T.

APPLICANT: Lowitz, Kevin P.

TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human

TITLE OR INVENTION: Receptors

FILE REFERENCE: AREN-0308

CURRENT APPLICATION NUMBER: US/09/995,225

CURRENT FILING DATE: 2001-11-26

PRIOR APPLICATION NUMBER: 09/170,496

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253,404

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: PCT/US99/23938

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253,404

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/255,366

PRIOR FILING DATE: 2000-12-12

PRIOR APPLICATION NUMBER: 60/270,286

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,365

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/270,266

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,032

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,358

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,356

PRIOR FILING DATE: 2001-04-06

```
/ PRIOR APPLICATION NUMBER: 60/290,917
/ PRIOR FILING DATE: 2001-05-14
/ PRIOR APPLICATION NUMBER: 60/309,208
/ PRIOR FILING DATE: 2001-07-31
/ NUMBER OF SEQ ID NOS: 67
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1002
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: No. US20020193584A1 Sequence
US-09-995-225-1
```

```
Query Match      100.0%; Score 1002; DB 9; Length 1002;
Best Local Similarity 100.0%; Pred. No. 7e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGTGAGCTGCAAGCTGTTCAACGCGCACAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60
Db 1 ATGTGAGCTGCAAGCTGTTCAACGCGCACAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60
QY 61 GACCTGAGCTGGGGCTGTCACTGTTGTGCTGCTGCGCTGTGGAGCTGTGGAGCTG 120
Db 61 GACCTGAGCTGGGGCTGTCACTGTTGTGCTGCTGCGCTGTGGAGCTGTGGAGCTG 120
QY 121 GGCCTGTGCTACACAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 GGCCTGTGCTACACAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 ATGCCGAGCTGTACTTGTGTCAACATGCGAGTGGCAGGCTGTGCTGCTGCTGCTG 240
Db 181 ATGCCGAGCTGTACTTGTGTCAACATGCGAGTGGCAGGCTGTGCTGCTGCTGCTG 240
QY 241 CCGTGTGACCTGCTGCGGCCCCCGAGAGCTCCCGGTGGCGCTGTGGAGTGTGGCGG 300
Db 241 CCGTGTGACCTGCTGCGGCCCCCGAGAGCTCCCGGTGGCGCTGTGGAGTGTGGCGG 300
QY 301 GTCCAGTGGCACTGCAAGATCCCTTCAATGTGCTCACTGCTGCTGCTGCTGCTGCT 360
Db 301 GTCCAGTGGCACTGCAAGATCCCTTCAATGTGCTCACTGCTGCTGCTGCTGCTGCT 360
QY 361 GGCCTGTGAGCTGCAACATGAGCGGTGCACTGCGCGGAGCTTACATGAGCGCAGC 420
Db 361 GGCCTGTGAGCTGCAACATGAGCGGTGCACTGCGCGGAGCTTACATGAGCGCAGC 420
QY 421 GTGTACAACAGCGGCGAGCTGTGCGGCTTCTGTGGGGTGGCGGCTGCTGACAGCTTC 480
Db 421 GTGTACAACAGCGGCGAGCTGTGCGGCTTCTGTGGGGTGGCGGCTGCTGACAGCTTC 480
QY 481 TCTCTGCTGCTCTTCTACATGTCAGCAGCATGTGTCCAGCCGCGGCTAAGTGGCGCAG 540
Db 481 TCTCTGCTGCTCTTCTACATGTCAGCAGCATGTGTCCAGCCGCGGCTAAGTGGCGCAG 540
QY 541 ATGCGAAGAGCAGAGCTGCCAGCGCGCAGCTGTGTTCATCGGCTACGTGGTGCAGCA 600
Db 541 ATGCGAAGAGCAGAGCTGCCAGCGCGCAGCTGTGTTCATCGGCTACGTGGTGCAGCA 600
QY 601 CTGGCCACCTCTACAGCGCTGTGTCTCTCCCGGTCCGAGGAGGAGCAGCGCCCTG 660
Db 601 CTGGCCACCTCTACAGCGCTGTGTCTCTCCCGGTCCGAGGAGGAGCAGCGCCCTG 660
QY 661 GACCGGAGCAGCGGCGGCTGTGAGCGCTGTGCGCAGCAGAGCTGTGGTGCACCGGTG 720
Db 661 GACCGGAGCAGCGGCGGCTGTGAGCGCTGTGCGCAGCAGAGCTGTGGTGCACCGGTG 720
QY 721 AGCGATTGGGCTGTGAGCGCACAATATGTGATCTGTGCTGGGAGCAGCGGTCAATC 780
Db 721 AGCGATTGGGCTGTGAGCGCACAATATGTGATCTGTGCTGGGAGCAGCGGTCAATC 780
QY 781 TGGCAGAGGAGAGCGCGTGTGAGCAGCACAATATGTGAGGAGCTTTGTGAGAGATTTC 840
Db 781 TGGCAGAGGAGAGCGCGTGTGAGCAGCACAATATGTGAGGAGCTTTGTGAGAGATTTC 840
```

```
QY 841 TCCAACTCTGCGCTTCTTCAGACAGCTTTGTGACACCACTTCTTACCGCTACATGAC 900
Db 841 TCCAACTCTGCGCTTCTTCAGACAGCTTTGTGACACCACTTCTTACCGCTACATGAC 900
QY 901 CAGAGCTTCCCGAGCAGAGCTTCCAAAGGCTGATGATAAAAGCTGCCGCGGAGCCGCGAC 960
Db 901 CAGAGCTTCCCGAGCAGAGCTTCCAAAGGCTGATGATAAAAGCTGCCGCGGAGCCGCGAC 960
QY 961 TGTCTCCCGAGCAGCAGTGGGGGTGACAGAGTGTGGCGTAG 1002
Db 961 TGTCTCCCGAGCAGCAGTGGGGGTGACAGAGTGTGGCGTAG 1002
```

RESULT 2

```
US-09-990-940-3
/ Sequence 3, Application US/0990940
/ Publication No. US20030027252A1
/ GENERAL INFORMATION:
/ APPLICANT: Tian, Hui
/ APPLICANT: Zhao, Jiayang
/ APPLICANT: Chen, Jin-Long
/ APPLICANT: Cutler, Gene
/ APPLICANT: An, Songzhu
/ APPLICANT: Dai, Kang
/ APPLICANT: Gupta, Jamila S.
/ APPLICANT: Thiarik Inc.
/ TITLE OF INVENTION: No. US20030027252A1 Receptors
/ FILE REFERENCE: 018781-007410US
/ CURRENT APPLICATION NUMBER: US/09/990,940
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/252,841
/ PRIOR FILING DATE: 2000-11-22
/ PRIOR APPLICATION NUMBER: US 60/257,636
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 60/261,377
/ PRIOR FILING DATE: 2001-01-12
/ PRIOR APPLICATION NUMBER: US 60/279,554
/ PRIOR FILING DATE: 2001-03-28
/ PRIOR APPLICATION NUMBER: US 60/280,696
/ PRIOR FILING DATE: 2001-03-29
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 1002
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1002)
/ OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR339
US-09-990-940-3
```

```
Query Match      100.0%; Score 1002; DB 10; Length 1002;
Best Local Similarity 100.0%; Pred. No. 7e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGTGAGCTGCAAGCTGTTCAACGCGCACAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60
Db 1 ATGTGAGCTGCAAGCTGTTCAACGCGCACAGGGCTGTGGAGAGCTGCTGCTGCGCAG 60
QY 61 GACCTGAGCTGGGGCTGTCACTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 GACCTGAGCTGGGGCTGTCACTGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 GGCCTGTGCTACACAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 121 GGCCTGTGCTACACAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 ATGCCGAGCTGTACTTGTGTCAACATGCGAGTGGCAGGCTGTGCTGCTGAGCGCTG 240
Db 181 ATGCCGAGCTGTACTTGTGTCAACATGCGAGTGGCAGGCTGTGCTGCTGAGCGCTG 240
```

```
QY 241 CCTGTGACCTGCTCGGAGCCCCCGAGAGTCCCGGGGCGCTGTGAGTGTGGGCGGAA 300
| | | | |
Db 241 CCTGTGACCTGCTCGGAGCCCCCGAGAGTCCCGGGGCGCTGTGAGTGTGGGCGGAA 300
QY 301 GTTCACGTGGCACTGAGATCCCTTCATATGTGTCTCACTGTGTGGCATGTATCTCACC 360
| | | | |
Db 301 GTTCACGTGGCACTGAGATCCCTTCATATGTGTCTCACTGTGTGGCATGTATCTCACC 360
QY 361 GGCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCTCCGGGAGCACTATGCGCAGC 420
| | | | |
Db 361 GGCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCTCCGGGAGCACTATGCGCAGC 420
QY 421 GTGTACAAACGCGGAGCGGTGTGAGGCTGTGTGGGGGTGGCGGCTGTGACCAAGTTTC 480
| | | | |
Db 421 GTGTACAAACGCGGAGCGGTGTGAGGCTGTGTGGGGGTGGCGGCTGTGACCAAGTTTC 480
QY 481 TCCTGCTGTCTTCTTACATCTGTGACGATGTGTCCACCCGCGCTAGAGTGGCCAG 540
| | | | |
Db 481 TCCTGCTGTCTTCTTACATCTGTGACGATGTGTCCACCCGCGCTAGAGTGGCCAG 540
QY 541 ATGCAGAAACGAGAGCTGCGACGCGCACTGTGTTCATCGGCTACGTTGTCAGCA 600
| | | | |
Db 541 ATGCAGAAACGAGAGCTGCGACGCGCACTGTGTTCATCGGCTACGTTGTCAGCA 600
QY 601 CTGGCAACCTCTACGCGGTGTGTACTCTCCGCGCTCCGAGGAGGAGCAGCCCTTG 660
| | | | |
Db 601 CTGGCAACCTCTACGCGGTGTGTACTCTCCGCGCTCCGAGGAGGAGCAGCCCTTG 660
QY 661 GACCGGAGACGCGGCGCGCTGTGAGCCCTCGGACACAGGCTGTGTGGCCACCGTGTGC 720
| | | | |
Db 661 GACCGGAGACGCGGCGCGCTGTGAGCCCTCGGACACAGGCTGTGTGGCCACCGTGTGC 720
QY 721 AGCGATTTGGGCTCTGTGACGCACTATCTGTGATCTGTGTGGGAGCAACGTTGATATC 780
| | | | |
Db 721 AGCGATTTGGGCTCTGTGACGCACTATCTGTGATCTGTGTGGGAGCAACGTTGATATC 780
QY 781 TCGCGAGGAGAGCGCGGTGTGACGCACTATCTGTGATCTGTGTGGGAGGATTC 840
| | | | |
Db 781 TCGCGAGGAGAGCGCGGTGTGACGCACTATCTGTGATCTGTGTGGGAGGATTC 840
QY 841 TCCAAATCTCTGCTCTTCTGTGACGATGTGTGACCACTTCTTACCGCTACATGAAAC 900
| | | | |
Db 841 TCCAAATCTCTGCTCTTCTGTGACGATGTGTGACCACTTCTTACCGCTACATGAAAC 900
QY 901 CAGAGCTTCCCGACGACGCTCCAAACGCGTGTATGAAAAAGCTGTGCGGAGACCGGAC 960
| | | | |
Db 901 CAGAGCTTCCCGACGACGCTCCAAACGCGTGTATGAAAAAGCTGTGCGGAGACCGGAC 960
QY 961 TGCTCCCGGACCACTATGGGGGTGTGACGAGTGTGCGGTAG 1002
| | | | |
Db 961 TGCTCCCGGACCACTATGGGGGTGTGACGAGTGTGCGGTAG 1002

RESULT 3
US-09-995-225-1
; Sequence 1, Application US/0995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
```

```
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9el Sequence
US-09-995-225-1
```

```
Query Match 100.0%; Score 1002; DB 10; Length 1002;
Best Local Similarity 100.0%; Pred. No. 76-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGTGAGCTGACGCTGTGTTCAACGAGCAGAGGCTGTGAGAGCTGTGCTGTGCAG 60
| | | | |
Db 1 ATGTGAGCTGACGCTGTGTTCAACGAGCAGAGGCTGTGAGAGCTGTGCTGTGCAG 60
QY 61 GACCTGACGCTGGGCGTGTACCTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 120
| | | | |
Db 61 GACCTGACGCTGGGCGTGTACCTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 120
QY 121 GGCCTGTGTTACAAACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
| | | | |
Db 121 GGCCTGTGTTACAAACGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 181 ATGCCGAGCGTGTACTTGTGTCAATGTGCAATGTGCAATGTGCAATGTGCAATGTGCAATGTG 240
| | | | |
Db 181 ATGCCGAGCGTGTACTTGTGTCAATGTGCAATGTGCAATGTGCAATGTGCAATGTGCAATGTG 240
QY 241 CCTGTGACCTGCTCGGCCCCCGAGAGCTCCCGGTGGCGCTGTGTGTGTGTGTGTGTGTGTGT 300
| | | | |
Db 241 CCTGTGACCTGCTCGGCCCCCGAGAGCTCCCGGTGGCGCTGTGTGTGTGTGTGTGTGTGTGT 300
QY 301 GTTCACGTGGCACTGAGATCCCTTCATATGTGTCTCACTGTGTGGCATGTATCTCACC 360
| | | | |
Db 301 GTTCACGTGGCACTGAGATCCCTTCATATGTGTCTCACTGTGTGGCATGTATCTCACC 360
QY 361 GGCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCTCCGGGAGCACTATGCGCAGC 420
| | | | |
Db 361 GGCCTGTGAGCTTCGACCACTATGAGCGTGTGACATGCTCCGGGAGCACTATGCGCAGC 420
QY 421 GTGTACAAACGCGGAGCGGTGTGAGGCTGTGTGGGGGTGGCGGCTGTGACCAAGTTTC 480
| | | | |
Db 421 GTGTACAAACGCGGAGCGGTGTGAGGCTGTGTGGGGGTGGCGGCTGTGACCAAGTTTC 480
QY 481 TCCTGCTGTCTTCTTACATCTGTGACGATGTGTCCACCCGCGCTAGAGTGGCCAG 540
| | | | |
Db 481 TCCTGCTGTCTTCTTACATCTGTGACGATGTGTCCACCCGCGCTAGAGTGGCCAG 540
QY 541 ATGCAGAAACGAGAGCTGCGACGCGCACTGTGTTCATCGGCTACGTTGTCAGCA 600
| | | | |
Db 541 ATGCAGAAACGAGAGCTGCGACGCGCACTGTGTTCATCGGCTACGTTGTCAGCA 600
```

QY 601 CTGGCCACCTCTACGCGCTGTGTCTACTCTCCCGGCTCCGACGGGAGGACACGCCCTTG 660
Db 601 CTGGCCACCTCTACGCGCTGTGTCTACTCTCCCGGCTCCGACGGGAGGACACGCCCTTG 660
QY 661 GACCGGGACACGCGCGCTGTGAGCCCTCGGACACAGAGCTGTGTGTGACCGCTGTG 720
Db 661 GACCGGGACACGCGCGCTGTGAGCCCTCGGACACAGAGCTGTGTGTGACCGCTGTG 720
QY 721 ACCGAGTTGGGCTCTGGGAGCGGACCACTATCTGATCTGTGGGGACACAGCTATCATC 780
Db 721 ACCGAGTTGGGCTCTGGGAGCGGACCACTATCTGATCTGTGGGGACACAGCTATCATC 780
QY 781 TCGCAGGGAAGCCCGTGTGACGACACTACCTGGGGCTACTGCACTTTGTGAAGATTTC 840
Db 781 TCGCAGGGAAGCCCGTGTGACGACACTACCTGGGGCTACTGCACTTTGTGAAGATTTC 840
QY 841 TCCAAACTCTGTGCTTCTTCAGAGCTTTGTGACACCACTTGTACCGCTACATGAAC 900
Db 841 TCCAAACTCTGTGCTTCTTCAGAGCTTTGTGACACCACTTGTACCGCTACATGAAC 900
QY 901 CAGAGCTTCCCGACCAAGCTCCAAAGCTGATGAAAAGCTGCCCGGGGACCGGAC 960
Db 901 CAGAGCTTCCCGACCAAGCTCCAAAGCTGATGAAAAGCTGCCCGGGGACCGGAC 960
QY 961 TGCTCCCGGACCAATGGGGGTGTGACAGAGTGTGCGGTAG 1002
Db 961 TGCTCCCGGACCAATGGGGGTGTGACAGAGTGTGCGGTAG 1002

RESULT 4
US-10-292-798-1015
Sequence 1015, Application US/10292798
Publication No. US2003023583A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: KATAYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1015
LENGTH: 1402
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE: (1) .. (1402)
NAME/KEY: CDS
LOCATION: (201) .. (1202)
US-10-292-798-1015

Query Match 100.0%; Score 1002; DB 15; Length 1402;
Best Local Similarity 100.0%; Pred. No. 7,3e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGTGAGCTGTTCACGACAGAGGCTGTGTGAGAGACTGTGCTGCTGCAG 60
Db 201 ATGTGAGCTGTGAGCTGTTCACGACAGAGGCTGTGTGAGAGACTGTGCTGCTGCAG 260
QY 61 GACCTGAGAGCTGTGCTGT 120
Db 261 GACCTGAGAGCTGTGCTGT 320

QY 121 GGCCTGTGCTAACAGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db 321 GGCCTGTGCTAACAGCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 380
QY 181 ATGCGGAGCTGTGCTGT 240
Db 381 ATGCGGAGCTGTGCTGT 440
QY 241 CCTGTGACCTGT 300
Db 441 CCTGTGACCTGT 500
QY 301 GTCCAGTGGGACCTGACATATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db 501 GTCCAGTGGGACCTGACATATCCCTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
QY 361 GCCCTGTGAGCTGTGACCACTACATGAGAGCTGTGACCTGTGACCTGTGACCTGTGAC 420
Db 561 GCCCTGTGAGCTGTGACCACTACATGAGAGCTGTGACCTGTGACCTGTGACCTGTGAC 620
QY 421 GTGTACAAACGCGGACAGT 480
Db 621 GTGTACAAACGCGGACAGT 680
QY 481 TCTGTGCTGT 540
Db 681 TCTGTGCTGT 740
QY 541 ATGCAAGAACGAGAGCTGTGACGACGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 741 ATGCAAGAACGAGAGCTGTGACGACGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 800
QY 601 CTGGCCACCTCTACGCGCTGTGTCTACTCTCCCGGCTCCGACGGGAGGACACGCCCTTG 660
Db 801 CTGGCCACCTCTACGCGCTGTGTCTACTCTCCCGGCTCCGACGGGAGGACACGCCCTTG 860
QY 661 GACCGGGACACGCGCGCTGTGAGCCCTCGGACACAGAGCTGTGTGTGTGTGTGTGTGT 720
Db 861 GACCGGGACACGCGCGCTGTGAGCCCTCGGACACAGAGCTGTGTGTGTGTGTGTGTGT 920
QY 721 ACGAGTTTGGGCTCTGTGACGCGCACATATCTGATCTGTGTGTGTGTGTGTGTGTGTGT 780
Db 921 ACGAGTTTGGGCTCTGTGACGCGCACATATCTGATCTGTGTGTGTGTGTGTGTGTGTGT 980
QY 781 TCGCAGGGAAGCCCGTGTGACGACACTACCTGGGGCTACTGCACTTTGTGAAGATTTC 840
Db 981 TCGCAGGGAAGCCCGTGTGACGACACTACCTGGGGCTACTGCACTTTGTGAAGATTTC 1040
QY 841 TCCAAACTCTGTGCTTCTTCAGAGCTTTGTGACACCACTTGTACCGCTACATGAAC 900
Db 1041 TCCAAACTCTGTGCTTCTTCAGAGCTTTGTGACACCACTTGTACCGCTACATGAAC 1100
QY 901 CAGAGCTTCCCGACCAAGCTCCAAAGCTGATGAAAAGCTGCCCGGGGACCGGAC 960
Db 1101 CAGAGCTTCCCGACCAAGCTCCAAAGCTGATGAAAAGCTGCCCGGGGACCGGAC 1160
QY 961 TGCTCCCGGACCAATGGGGGTGTGACAGAGTGTGCGGTAG 1002
Db 1161 TGCTCCCGGACCAATGGGGGTGTGACAGAGTGTGCGGTAG 1202

RESULT 5
US-10-017-161-1201
Sequence 1201, Application US/10017161
Publication No. US2003014368A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: KATAYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161

```

; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1201
; LENGTH: 1561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1561)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1361)
; US-10-017-161-1201

```

```

Query Match      100.0%; Score 1002; DB 15; Length 1561;
Best Local Similarity 100.0%; Pred. No. 7,4e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATGTGAGCTGCACTGGTTCACACGCGACAGGGCTGTGAGAGAGCTGCTGCTGCGAG 60
DB      360 ATGTGAGCTGCACTGGTTCACACGCGACAGGGCTGTGAGAGAGCTGCTGCTGCGAG 419
QY      61 GACCTGCACTGGGCGCTGTCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB      420 GACCTGCACTGGGCGCTGTCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479
QY      121 GGGCTGTGCTACCAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB      480 GGGCTGTGCTACCAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
QY      181 ATGCGGACGCTGTACTTTTTCACATGAGCAGTGGCAGGCTGTGCTGCTGCTGCTGCTG 240
DB      540 ATGCGGACGCTGTACTTTTTCACATGAGCAGTGGCAGGCTGTGCTGCTGCTGCTGCTG 599
QY      241 CCGTGTGCACTGCTGCGGCCCCCGAGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAG 300
DB      600 CCGTGTGCACTGCTGCGGCCCCCGAGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAG 659
QY      301 GTCACAGTGGCAGTGGCAGATCCCTTCAATGTGCTCCTGCTGCTGCTGCTGCTGCTGCTG 360
DB      660 GTCACAGTGGCAGTGGCAGATCCCTTCAATGTGCTCCTGCTGCTGCTGCTGCTGCTGCTG 719
QY      361 GCGCTGTGAGCTGCACTGACATGAGAGCTGCACTGCGCGGAGCCTTACATGAGCGAGC 420
DB      720 GCGCTGTGAGCTGCACTGACATGAGAGCTGCACTGCGCGGAGCCTTACATGAGCGAGC 779
QY      421 GTGTACACACGCGGCGCAGTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB      780 GTGTACACACGCGGCGCAGTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 839
QY      481 TCTGTGCTGCTTCTTCACTGAGCAGTGTGCACTGCGCGGAGCCTTACATGAGCGAGC 540
DB      840 TCTGTGCTGCTTCTTCACTGAGCAGTGTGCACTGCGCGGAGCCTTACATGAGCGAGC 899
QY      541 ATGCAAGACGAGAGCTGCGAGCGCAGCAGCTGTGTTCAATCGCTACATGCTGTGCGCAG 600
DB      900 ATGCAAGACGAGAGCTGCGAGCGCAGCAGCTGTGTTCAATCGCTACATGCTGTGCGCAG 959
QY      601 CTGGCCACCTCTTACAGCGCTGTGCTACTTCCCGGCTCGGACGAGAGACAGCGCCCTG 660
DB      960 CTGGCCACCTCTTACAGCGCTGTGCTACTTCCCGGCTCGGACGAGAGACAGCGCCCTG 1019
QY      661 GACCGGAGACGCGGCGGCTGAGGCGCTGCGGACACAGGCTGTGAGGAGACAGCGCTGTC 720
DB      1020 GACCGGAGACGCGGCGGCTGAGGCGCTGCGGACACAGGCTGTGAGGAGACAGCGCTGTC 1079
QY      721 ACGCAGTTTGGGCTGTGAGACGCACTATCTGATCTGTGCTGTGAGGACACAGCTATCATC 780
DB      1080 ACGCAGTTTGGGCTGTGAGACGCACTATCTGATCTGTGCTGTGAGGAGACAGCTATCATC 1139

```

```

QY      781 TCGCGAGGAGAGCGCGTGTGACGCACTACTGTGGGCTACTGTGCACTTTGTGAGAGATTTC 840
DB      1140 TCGCGAGGAGAGCGCGTGTGACGCACTACTGTGGGCTACTGTGCACTTTGTGAGAGATTTC 1199
QY      841 TCCAAATCCTGTGCTTCTTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 900
DB      1200 TCCAAATCCTGTGCTTCTTCCAGCAGCTTTGTGACACCACTTCTTACCGCTACATGAAC 1259
QY      901 CAGAGCTTCCCGACGCAAGCTTCCCAAGCGCTGTATGAAAAAAGTCCCTCGCGGAGCCGCGAC 960
DB      1260 CAGAGCTTCCCGACGCAAGCTTCCCAAGCGCTGTATGAAAAAAGTCCCTCGCGGAGCCGCGAC 1319
QY      961 TGCTCCCGGAGACCATGTGGGCGGTGACAGAGTGTGCGCTGAG 1002
DB      1320 TGCTCCCGGAGACCATGTGGGCGGTGACAGAGTGTGCGCTGAG 1361

```

RESULT 6

```

US-10-101-148-2
; Sequence 2, Application US/10101148
; Publication No. US20030187198A1
; GENERAL INFORMATION:
; APPLICANT: Yeung, George
; TITLE OF INVENTION: A No. US20030187198A1el Chemokine Receptor Obtained From a cDNA
; FILE REFERENCE: 28110/38317
; CURRENT APPLICATION NUMBER: US/10/101,148
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-101-148-2

```

```

Query Match      100.0%; Score 1002; DB 15; Length 1614;
Best Local Similarity 100.0%; Pred. No. 7,4e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATGTGAGCTGCACTGGTTCACACGCGACAGGGCTGTGAGAGAGCTGCTGCTGCGAG 60
DB      129 ATGTGAGCTGCACTGGTTCACACGCGACAGGGCTGTGAGAGAGCTGCTGCTGCGAG 188
QY      61 GACCTGCACTGGGCGCTGTCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB      189 GACCTGCACTGGGCGCTGTCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 248
QY      121 GGGCTGTGCTACCAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB      249 GGGCTGTGCTACCAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308
QY      249 GGGCTGTGCTACCAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308
DB      249 GGGCTGTGCTACCAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308
QY      181 ATGCGGACGCTGTACTTTTTCACATGAGCAGTGGCAGGCTGTGCTGCTGCTGCTGCTG 240
DB      309 ATGCGGACGCTGTACTTTTTCACATGAGCAGTGGCAGGCTGTGCTGCTGCTGCTGCTG 368
QY      241 CCGTGTGCACTGCTGCGGCCCCCGAGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAG 300
DB      369 CCGTGTGCACTGCTGCGGCCCCCGAGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAG 428
QY      301 GTCACAGTGGCAGTGGCAGATCCCTTCAATGTGCTCCTGCTGCTGCTGCTGCTGCTGCTG 360
DB      429 GTCACAGTGGCAGTGGCAGATCCCTTCAATGTGCTCCTGCTGCTGCTGCTGCTGCTGCTG 488
QY      429 GTCACAGTGGCAGTGGCAGATCCCTTCAATGTGCTCCTGCTGCTGCTGCTGCTGCTGCTG 488
DB      429 GTCACAGTGGCAGTGGCAGATCCCTTCAATGTGCTCCTGCTGCTGCTGCTGCTGCTGCTG 488
QY      361 GCGCTGTGAGCTGCACTGACATGAGAGCTGCACTGCGCGGAGCCTTACATGAGCGAGC 420
DB      489 GCGCTGTGAGCTGCACTGACATGAGAGCTGCACTGCGCGGAGCCTTACATGAGCGAGC 548
QY      421 GTGTACACACGCGGCGCAGTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB      549 GTGTACACACGCGGCGCAGTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 608

```

481 TCCTGCTGCTCTTACATCTGACGCCATGTTCCACCCGCGCTAGAGTCCCAAG 540
Db TCCTGCTGCTCTTACATCTGACGCCATGTTCCACCCGCGCTAGAGTCCCAAG 668
QY 541 ATGCAAGAGCAAGAGCTGCGCAGCGCAGCTGTTTATCGGCTAAGTGGTGCAGCA 600
Db 669 ATGCAAGAGCAAGAGCTGCGCAGCGCAGCTGTTTATCGGCTAAGTGGTGCAGCA 728
QY 601 CTGGCCACCTCTTACGCGCTGTTGCTACTCTCCCGCGCTCCGACAGGAGAGCAACCGCCCTG 660
Db 729 CTGGCCACCTCTTACGCGCTGTTGCTACTCTCCCGCGCTCCGACAGGAGAGCAACCGCCCTG 788
QY 661 GACCGGAGACAGGCGCGCTGAGAGCCCTCGGACACAGGCTGTTGAGCCAGCGTATGC 720
Db 789 GACCGGAGACAGGCGCGCTGAGAGCCCTCGGACACAGGCTGTTGAGCCAGCGTATGC 848
QY 721 AGCAGTTTGGGCTCTGAGCCGCACTATCTGATCTCTGCTGGGAGCAACGGTATCATC 780
Db 849 AGCAGTTTGGGCTCTGAGCCGCACTATCTGATCTCTGCTGGGAGCAACGGTATCATC 908
QY 781 TCGCGAGGGAAGCCGCTGAGAGCACTACTGAGGAGCTGACCTTGTGAGAGATTC 840
Db 909 TCGCGAGGGAAGCCGCTGAGAGCACTACTGAGGAGCTGACCTTGTGAGAGATTC 968
QY 841 TCCAACTCTCTGCTGCTTCTCCAGAGCTTGTGACACCACTTCTTACCGCTACATGAC 900
Db 969 TCCAACTCTCTGCTGCTTCTCCAGAGCTTGTGACACCACTTCTTACCGCTACATGAC 1028
QY 901 CAGAGCTTCCCAAGCAAGCTTCCAAAGCTGATGAGAAAGAGCGCCGCGGAGACCGGAC 960
Db 1029 CAGAGCTTCCCAAGCAAGCTTCCAAAGCTGATGAGAAAGAGCGCCGCGGAGACCGGAC 1088
QY 961 TGTCTCCCGGACCAATGAGGAGTGCAGAGTGTGCGGCTAG 1002
Db 1089 TGTCTCCCGGACCAATGAGGAGTGCAGAGTGTGCGGCTAG 1130

RESULT 7
US-09-791-932-54
; Sequence 54, Application US/09/91932
; Publication No. US2003003451A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; APPLICANT: Parodi, Luis A.
; APPLICANT: Hiesch, Ronald R.
; APPLICANT: Lind, Peter
; APPLICANT: Kayles, Paul S.
; APPLICANT: Ruff, Valerie
; APPLICANT: Huff, Rita M.
; APPLICANT: Wood, Linda S.
; TITLE OF INVENTION: No. US2003003451A1 G Protein-Coupled Receptors Cross-Referent
; FILE REFERENCE: 00325, US1
; CURRENT APPLICATION NUMBER: US/09/791, 932
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/184,305
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,304
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,303
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,397
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,247
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/188,880
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/217,369
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/217,370
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/218,492
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/186,810

; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/188,064
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: 60/186,457
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: 60/213,861
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/194,344
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 60/218,337
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-791-932-54

Query Match 100.0%; Score 1002; DB 10; Length 1778;
Best Local Similarity 100.0%; Pred. No. 7, 5e-282;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGACAGCTGTTTCAAGGCAAGGCTGTGAGAGCTGCTGCTGCAAG 60
Db 19 ATGTGAGCTGACAGCTGTTTCAAGGCAAGGCTGTGAGAGCTGCTGCTGCAAG 78
QY 61 GACCTGAGCTGAGGCTGTGCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 79 GACCTGAGCTGAGGCTGTGCACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 138
QY 121 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 139 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 198
QY 181 ATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 199 ATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 258
QY 241 CCTGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 259 CCTGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 318
QY 301 GTCCAGTGTGACATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 319 GTCCAGTGTGACATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378
QY 361 GCCCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 379 GCCCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 438
QY 421 GTGTACAAACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 439 GTGTACAAACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
QY 481 TCCTGCTGCTCTTACATCTGACGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 499 TCCTGCTGCTCTTACATCTGACGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 558
QY 541 ATGCAAGAGCAAGAGCTGCGCAGCGCAGCTGTTTATCGGCTAAGTGGTGCAGCA 600
Db 559 ATGCAAGAGCAAGAGCTGCGCAGCGCAGCTGTTTATCGGCTAAGTGGTGCAGCA 618
QY 601 CTGGCCACCTCTTACGCGCTGTTGCTACTCTCCCGCGCTCCGACAGGAGAGCAACCGCCCTG 660
Db 619 CTGGCCACCTCTTACGCGCTGTTGCTACTCTCCCGCGCTCCGACAGGAGAGCAACCGCCCTG 678
QY 661 GACCGGAGACAGGCGCGCTGAGAGCCCTCGGACACAGGCTGTTGAGCCAGCGTATGC 720
Db 679 GACCGGAGACAGGCGCGCTGAGAGCCCTCGGACACAGGCTGTTGAGCCAGCGTATGC 738
QY 721 AGCAGTTTGGGCTCTGAGCCGCACTATCTGATCTCTGCTGGGAGCAACGGTATCATC 780

Db 739 ACCGAGTTTGGGCTGTGAGACGCCACACTATCTGATCTGCTGGGGGACACAGGTATCATC 798
Qy 781 TCCGAGGAGAAAGCCCGGTGAGACGACACTACTGAGGGGCTACTGCACTTTGTGAAGATTTC 840
Db 799 TCCGAGGAGAAAGCCCGGTGAGACGACACTACTGAGGGGCTACTGCACTTTGTGAAGATTTC 858
Qy 841 TCCAACTCTGAGCTTTCTCAGAGAGTTGTGACACACTTCTCTACCGGCTACATGAAC 900
Db 859 TCCAACTCTGAGCTTTCTCAGAGAGTTGTGACACACTTCTCTACCGGCTACATGAAC 918
Qy 901 CAGAGCTTCCCGACGACGCTCCAAACGGCTGATGAAAAAGCTGCTGCGGGGACCGGAC 960
Db 919 CAGAGCTTCCCGACGACGCTCCAAACGGCTGATGAAAAAGCTGCTGCGGGGACCGGAC 978
Qy 961 TGTCTCCCGGACCAATGAGGGGTGACAGAGGTGCTGCGGTAG 1002
Db 979 TGTCTCCCGGACCAATGAGGGGTGACAGAGGTGCTGCGGTAG 1020

RESULT 8
US-10-225-567A-679
; Sequence 679, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 679
; LENGTH: 1785
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-679

Query Match 100.0%; Score 1002; DB 15; Length 1785;
Best Local Similarity 100.0%; Pred. No. 7.5e-262;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGAGCTGACAGCTGTTTCAACGACGACGAGCTGTGAGAGAGCTGCTGCTGCGAC 60
Db 21 ATGTGAGCTGACAGCTGTTTCAACGACGACGAGCTGTGAGAGAGCTGCTGCTGCGAC 80
Qy 61 GACCTGACGCTGGGGCTGTCTACTGTTGTGCTGCTGCGGCTGTGTTGGGCGTGCACATG 120
Db 81 GACCTGACGCTGGGGCTGTCTACTGTTGTGCTGCTGCGGCTGTGTTGGGCGTGCACATG 140
Qy 121 GGGCTGTGCAACAACCGCTGTGCTGCTGCGGCACTTAACAAGGACGACACATGAC 180
Db 141 GGGCTGTGCAACAACCGCTGTGCTGCTGCGGCACTTAACAAGGACGACACATGAC 200
Qy 181 ATGCGGAGCGTACTTGTTCATCATGAGAGAGCTGAGTGTGCTCAGCGCTGAGCC 240
Db 201 ATGCGGAGCGTACTTGTTCATCATGAGAGAGCTGAGTGTGCTCAGCGCTGAGCC 260
Qy 241 CCTGTGCACTGTGCGGCGGCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGAGAA 300
Db 261 CCTGTGCACTGTGCGGCGGCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGCGGAGAA 320
Qy 301 GTTCAAGTGGCACTGAGAGATCCGCTTCAATGATGCTTCACTGATGAGGCACTGATCCACC 360
Db 321 GTTCAAGTGGCACTGAGAGATCCGCTTCAATGATGCTTCACTGATGAGGCACTGATCCACC 380
Qy 361 GCCCTGTGAGCTGTGACCACTAATGAGAGCTGACATGCGGGGAGACTTACATGAGCCAGC 420
Db 381 GCCCTGTGAGCTGTGACCACTAATGAGAGCTGACATGCGGGGAGACTTACATGAGCCAGC 440

Qy 421 GTTACAAACACGCGGAGACGCTGTGCGGCTTCTGCTGCGGGGTGCGCGGCTGCTGACAGCTTC 480
Db 441 GTTACAAACACGCGGAGACGCTGTGCGGCTTCTGCTGCGGGGTGCGCGGCTGCTGACAGCTTC 500
Qy 481 TCTCTGCTGCTCTTCTACATCTGACAGCACTGATGCTCCAGCGGCTAGAGTGGCCAG 540
Db 501 TCTCTGCTGCTCTTCTACATCTGACAGCACTGATGCTCCAGCGGCTAGAGTGGCCAG 560
Qy 541 ATGCAAGACGACAGAGCTGCGGACGCGCAAGCTGCTGCTTCACTGCGCTTACGCTGACGAC 600
Db 561 ATGCAAGACGACAGAGCTGCGGACGCGCAAGCTGCTGCTTCACTGCGCTTACGCTGACGAC 620
Qy 601 CTGGCCACCTCTTACAGCGCTGTGCTACTCTCCCGGCTCCGAGGAGAGACAGCCCTTG 660
Db 621 CTGGCCACCTCTTACAGCGCTGTGCTACTCTCCCGGCTCCGAGGAGAGACAGCCCTTG 680
Qy 661 GACCGGAGACAGCGGCGGCTGTGAGAGCTTCCGACACAGAGCTGCTGCTGCGGACCGTGTGC 720
Db 681 GACCGGAGACAGCGGCGGCTGTGAGAGCTTCCGACACAGAGCTGCTGCTGCGGACCGTGTGC 740
Qy 721 AGCGAGTTTGGGCTGTGAGAGCGGACACTATCTGATCTGCTGCGGAGACAGGTATCATC 780
Db 741 AGCGAGTTTGGGCTGTGAGAGCGGACACTATCTGATCTGCTGCGGAGACAGGTATCATC 800
Qy 781 TCCGAGGAGAAAGCCCGGTGAGACGACACTACTGAGGGGCTACTGCACTTTGTGAAGATTTC 840
Db 801 TCCGAGGAGAAAGCCCGGTGAGACGACACTACTGAGGGGCTACTGCACTTTGTGAAGATTTC 860
Qy 841 TCCAACTCTGAGCTTTCTCAGAGAGCTTGTGACACCACTTCTTACCGCTACATGAAC 900
Db 861 TCCAACTCTGAGCTTTCTCAGAGAGCTTGTGACACCACTTCTTACCGCTACATGAAC 920
Qy 901 CAGAGCTTCCCGACGACGCTCCAAACGGCTGATGAAAAAGCTGCTGCGGGGACCGGAC 960
Db 921 CAGAGCTTCCCGACGACGCTCCAAACGGCTGATGAAAAAGCTGCTGCGGGGACCGGAC 980
Qy 961 TGTCTCCCGGACCAATGAGGGGTGACAGAGGTGCTGCGGTAG 1002
Db 981 TGTCTCCCGGACCAATGAGGGGTGACAGAGGTGCTGCGGTAG 1022

RESULT 9
US-10-165-844-14
; Sequence 14, Application US/10165844
; Publication No. US20030017539A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Hunter, John U.
; APPLICANT: Rudolph-Owen, Laura
; APPLICANT: Welch, Nadine S.
; APPLICANT: Slios-Santiago, Immaculada
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Encoding
; FILE REFERENCE: 35800/248302
; CURRENT APPLICATION NUMBER: US/10/165,844
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 09/088,857
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: US 09/324,465
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 09/464,685
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: US 09/741,783
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 09/145,745
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: US 09/383,745
; PRIOR FILING DATE: 1999-08-26
; PRIOR APPLICATION NUMBER: US 09/234,923
; PRIOR FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: US 09/340,880

; PRIOR FILING DATE: 1999-06-28
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: FaSTSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 1928
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (160)...(1158)
 US-10-165-844-14

Query Match 100.0%; Score 1002; DB 14; Length 1928;
 Best Local Similarity 100.0%; Pred. No. 7,6e-262;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGCAGCTGTTTCAACGGCAAGGCTGTGTGAGAGAGTCTGCTGCTGCCAG 60
 DB 160 ATGTGAGCTGCAGCTGTTTCAACGGCAAGGCTGTGTGAGAGAGTCTGCTGCTGCCAG 219
 QY 61 GACCTGCACTGGGGCTGTCACTGTTGTGCTGCTGGGCTGGTGGGCTGGCCAGTG 120
 DB 220 GACCTGCACTGGGGCTGTCACTGTTGTGCTGCTGGGCTGGTGGGCTGGCCAGTG 279
 QY 121 GGCCTGTGCTACAAAGCCCTGTGTGTGCTGGCCAACTTACACAGAAAGCCAGATGACC 180
 DB 280 GGCCTGTGCTACAAAGCCCTGTGTGTGCTGGCCAACTTACACAGAAAGCCAGATGACC 339
 QY 181 ATGCCGGAAGCTGTACTTTTTCACATGGGAGTGGGAGGCTGGTGGTCAAGGCTGGGAGC 240
 DB 340 ATGCCGGAAGCTGTACTTTTTCACATGGGAGTGGGAGGCTGGTGGTCAAGGCTGGGAGC 399
 QY 241 CCTGTGACCTGTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 300
 DB 400 CCTGTGACCTGTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 459
 QY 301 GTCCAGTGTGCACTGCAATCCCTTAAATGTGCTCTCATCTGGTGGCCATGTACTCCACC 360
 DB 460 GTCCAGTGTGCACTGCAATCCCTTAAATGTGCTCTCATCTGGTGGCCATGTACTCCACC 519
 QY 361 GGCCTGTGAGGCTGTGCACTGCAATCCCTTAAATGTGCTCTCATCTGGTGGCCATGTACTCCACC 420
 DB 520 GGCCTGTGAGGCTGTGCACTGCAATCCCTTAAATGTGCTCTCATCTGGTGGCCATGTACTCCACC 579
 QY 421 GTGTACAAACAGCGGCAAGTGTGCGGCTGTGTGTGGGCTGGGCTGGGCTGGGCTGGG 480
 DB 580 GTGTACAAACAGCGGCAAGTGTGCGGCTGTGTGTGGGCTGGGCTGGGCTGGGCTGGG 639
 QY 481 TCTCTGCTGCTCTTCTTACATCTGCAAGCAATGTGTCCACCGCGGCTGTAAGTGGCCAG 540
 DB 640 TCTCTGCTGCTCTTCTTACATCTGCAAGCAATGTGTCCACCGCGGCTGTAAGTGGCCAG 599
 QY 640 TCTCTGCTGCTCTTCTTACATCTGCAAGCAATGTGTCCACCGCGGCTGTAAGTGGCCAG 699
 DB 541 ATGCAAGAGCAAGAGCTGCGAGCGCAAGCTGTGTTCATCGGCTTACGTGTGCCAGCA 600
 DB 700 ATGCAAGAGCAAGAGCTGCGAGCGCAAGCTGTGTTCATCGGCTTACGTGTGCCAGCA 759
 QY 601 CTGGGCACTCTTACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 DB 760 CTGGGCACTCTTACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 819
 QY 661 GACCGGGAACAGGGGCTGGAGGCTTGGGCAACAGGCTGTGTGTGTGTGTGTGTGTGTGTGT 720
 DB 820 GACCGGGAACAGGGGCTGGAGGCTTGGGCAACAGGCTGTGTGTGTGTGTGTGTGTGTGTGT 879
 QY 721 ACGCAGTTTGGGCTTGTGACGCGCACATATCTGATCTGTGTGGGCAACGCGTATCATTC 780
 DB 880 ACGCAGTTTGGGCTTGTGACGCGCACATATCTGATCTGTGTGGGCAACGCGTATCATTC 939
 QY 781 TGGCAGAGGAGACCCGTGTGAGCGCACATCTGCGGGCTACTGCACTTTGTGAAGATTTTC 840
 DB 940 TGGCAGAGGAGACCCGTGTGAGCGCACATCTGCGGGCTACTGCACTTTGTGAAGATTTTC 999
 QY 841 TCCAAACTCTGCGCTTCTCCAGCAGCTTTGTGACACCACTTCTCTACCGCTACATGAAC 900

DB 1000 TCCAAACTCTGCGCTTCTCCAGCAGCTTTGTGACACCACTTCTCTACCGCTACATGAAC 1059
 QY 901 CAGAGCTTCCCGACAGAGCTTCAACGCGCTGATGATAAAAGCTGCTGGCGGACCGGAC 960
 DB 1060 CAGAGCTTCCCGACAGAGCTTCAACGCGCTGATGATAAAAGCTGCTGGCGGACCGGAC 1119
 QY 961 TGTCTCCCGGACCAATGGGGGTGCAAGCAAGTGTGCTGGCGCTAG 1002
 DB 1120 TGTCTCCCGGACCAATGGGGGTGCAAGCAAGTGTGCTGGCGCTAG 1161

RESULT 10
 US-09-791-932-31
 ; Sequence 31, Application US/09791932
 ; Publication No. US20030003451A1

; GENERAL INFORMATION:
 ; APPLICANT: Vogel, Gabriel
 ; APPLICANT: Parodi, Luis A.
 ; APPLICANT: Hiesbach, Ronald R.
 ; APPLICANT: Lind, Peter
 ; APPLICANT: Kaytee, Paul S.
 ; APPLICANT: Ruff, Valerie
 ; APPLICANT: Huff, Rita M.
 ; APPLICANT: Wood, Linda S.
 ; TITLE OF INVENTION: No. US20030003451A1el G Protein-Coupled Receptors Cross-Referen
 ; FILE REFERENCE: 00325.US1
 ; CURRENT APPLICATION NUMBER: US/09/791,932
 ; PRIOR FILING DATE: 2001-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,305
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,304
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,303
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,397
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/184,247
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/188,880
 ; PRIOR FILING DATE: 2000-03-13
 ; PRIOR APPLICATION NUMBER: 60/217,369
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 60/217,370
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 60/218,492
 ; PRIOR FILING DATE: 2000-07-20
 ; PRIOR APPLICATION NUMBER: 60/186,810
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/188,064
 ; PRIOR FILING DATE: 2000-03-09
 ; PRIOR APPLICATION NUMBER: 60/186,457
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: 60/213,861
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/194,344
 ; PRIOR FILING DATE: 2000-04-03
 ; PRIOR APPLICATION NUMBER: 60/218,337
 ; PRIOR FILING DATE: 2000-07-14
 ; NUMBER OF SEQ ID NOS: 184
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 31
 ; LENGTH: 1750
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

Query Match 99.8%; Score 1000.4; DB 10; Length 1750;
 Best Local Similarity 99.9%; Pred. No. 2e-261; 1; Indels 0; Gaps 0;
 Matches 1001; Conservative 0; Mismatches 1;

QY 1 ATGTGAGCTGCAGCTGTTTCAACGGCAAGGCTGTGTGAGAGAGTCTGCTGCTGCCAG 60

	558	GTGTACAACA	CAGCGG	GACG	TGTCGG	CTTTCG	TGTGGG	TGGG	GGCG	CTGTAC	CAAG	CTTC	617
Db	558	GTGTACAACA	CAGCGG <td>GACG<td>TGTCGG<td>CTTTCG<td>TGTGGG<td>TGGG<td>GGCG<td>CTGTAC<td>CAAG<td>CTTC</td><td>617</td></td></td></td></td></td></td></td></td>	GACG <td>TGTCGG<td>CTTTCG<td>TGTGGG<td>TGGG<td>GGCG<td>CTGTAC<td>CAAG<td>CTTC</td><td>617</td></td></td></td></td></td></td></td>	TGTCGG <td>CTTTCG<td>TGTGGG<td>TGGG<td>GGCG<td>CTGTAC<td>CAAG<td>CTTC</td><td>617</td></td></td></td></td></td></td>	CTTTCG <td>TGTGGG<td>TGGG<td>GGCG<td>CTGTAC<td>CAAG<td>CTTC</td><td>617</td></td></td></td></td></td>	TGTGGG <td>TGGG<td>GGCG<td>CTGTAC<td>CAAG<td>CTTC</td><td>617</td></td></td></td></td>	TGGG <td>GGCG<td>CTGTAC<td>CAAG<td>CTTC</td><td>617</td></td></td></td>	GGCG <td>CTGTAC<td>CAAG<td>CTTC</td><td>617</td></td></td>	CTGTAC <td>CAAG<td>CTTC</td><td>617</td></td>	CAAG <td>CTTC</td> <td>617</td>	CTTC	617
QY	481	TCCTCGTGTCTT	TTTACATCTG <td>CAAGCA</td> <td>TGTCGTCA</td> <td>CCCCCGG</td> <td>CTTAA</td> <td>GA</td> <td>TGTG</td> <td>CGCAAG</td> <td>540</td> <td></td> <td></td>	CAAGCA	TGTCGTCA	CCCCCGG	CTTAA	GA	TGTG	CGCAAG	540		
Db	618	TCCTCGTGTCTT	TTTAACTG <td>CAAGCA</td> <td>TGTGTCA</td> <td>CCCCCGG</td> <td>CTTAA</td> <td>GA</td> <td>TGTG</td> <td>CGCAAG</td> <td>677</td> <td></td> <td></td>	CAAGCA	TGTGTCA	CCCCCGG	CTTAA	GA	TGTG	CGCAAG	677		
QY	541	ATGCAGAACG <td>CAGAGCTT</td> <td>CGACG<td>CCACG<td>CTGTGTTCAT</td><td>CGGCTAC<td>AGTGTG<td>CCAGCA</td><td>600</td><td></td><td></td><td></td></td></td></td></td>	CAGAGCTT	CGACG <td>CCACG<td>CTGTGTTCAT</td><td>CGGCTAC<td>AGTGTG<td>CCAGCA</td><td>600</td><td></td><td></td><td></td></td></td></td>	CCACG <td>CTGTGTTCAT</td> <td>CGGCTAC<td>AGTGTG<td>CCAGCA</td><td>600</td><td></td><td></td><td></td></td></td>	CTGTGTTCAT	CGGCTAC <td>AGTGTG<td>CCAGCA</td><td>600</td><td></td><td></td><td></td></td>	AGTGTG <td>CCAGCA</td> <td>600</td> <td></td> <td></td> <td></td>	CCAGCA	600			
Db	678	ATGCAGAACG <td>CAGAGCTT</td> <td>CGACG<td>CCACG<td>CTGTGTTCAT</td><td>CGGCTAC<td>AGTGTG<td>CCAGCA</td><td>737</td><td></td><td></td><td></td></td></td></td></td>	CAGAGCTT	CGACG <td>CCACG<td>CTGTGTTCAT</td><td>CGGCTAC<td>AGTGTG<td>CCAGCA</td><td>737</td><td></td><td></td><td></td></td></td></td>	CCACG <td>CTGTGTTCAT</td> <td>CGGCTAC<td>AGTGTG<td>CCAGCA</td><td>737</td><td></td><td></td><td></td></td></td>	CTGTGTTCAT	CGGCTAC <td>AGTGTG<td>CCAGCA</td><td>737</td><td></td><td></td><td></td></td>	AGTGTG <td>CCAGCA</td> <td>737</td> <td></td> <td></td> <td></td>	CCAGCA	737			
QY	601	CTGGCAACCT	CTAAGCG <td>TGTGTCTA</td> <td>TTTCCG</td> <td>CGCTCG<td>CAAGGAG<td>ACA</td><td>CGCCCTTG</td><td>660</td><td></td><td></td><td></td></td></td>	TGTGTCTA	TTTCCG	CGCTCG <td>CAAGGAG<td>ACA</td><td>CGCCCTTG</td><td>660</td><td></td><td></td><td></td></td>	CAAGGAG <td>ACA</td> <td>CGCCCTTG</td> <td>660</td> <td></td> <td></td> <td></td>	ACA	CGCCCTTG	660			
Db	738	CTGGCAACCT	CTAAGCG <td>TGTGTCTA</td> <td>TTTCCG</td> <td>CGCTCG<td>CAAGGAG<td>ACA</td><td>CGCCCTTG</td><td>797</td><td></td><td></td><td></td></td></td>	TGTGTCTA	TTTCCG	CGCTCG <td>CAAGGAG<td>ACA</td><td>CGCCCTTG</td><td>797</td><td></td><td></td><td></td></td>	CAAGGAG <td>ACA</td> <td>CGCCCTTG</td> <td>797</td> <td></td> <td></td> <td></td>	ACA	CGCCCTTG	797			
QY	661	GACCGGAA	CAAGG <td>CGCTG<td>GAAGC<td>CTTCG<td>CAACAG</td><td>AGCTGTG<td>TGAGGCA</td><td>CCGTGTG</td><td>720</td><td></td><td></td></td></td></td></td>	CGCTG <td>GAAGC<td>CTTCG<td>CAACAG</td><td>AGCTGTG<td>TGAGGCA</td><td>CCGTGTG</td><td>720</td><td></td><td></td></td></td></td>	GAAGC <td>CTTCG<td>CAACAG</td><td>AGCTGTG<td>TGAGGCA</td><td>CCGTGTG</td><td>720</td><td></td><td></td></td></td>	CTTCG <td>CAACAG</td> <td>AGCTGTG<td>TGAGGCA</td><td>CCGTGTG</td><td>720</td><td></td><td></td></td>	CAACAG	AGCTGTG <td>TGAGGCA</td> <td>CCGTGTG</td> <td>720</td> <td></td> <td></td>	TGAGGCA	CCGTGTG	720		

3

QY	722	AGCAGTTTGGGCTTGTGAACCCACAATACTATCTGATCTCGTGGGGCAACAGCATCATC	780
Db	858	AACGAAATTGGGCTGTGGACGCCAACACTATCTTAATCTGCTGGGGGCACAGCATCATC	917
QY	781	TGCGAAGGGAAGCCCGTGAGCACCACTACCTGGGGCTACTGCACCTTGTGAAGATTTC	840
Db	918	TGCGAAGGGAAGCCCGTGAGCACCACTACCTGGGGCTACTGCACCTTGTGAAGATTTC	977
QY	841	TCCAAACCTCTGGGCTTCTCAGAGAGTTTGTGACACCACTTCTTAACGGCTACATGAC	900
Db	978	TCCAAAACCTCTGGGCTTCTCAGAGAGTTTGTGACACCACTTCTTAACGGCTACATGAC	1037
QY	901	CAGAAGCTTCCCACAGAAGCTCCAACGGCTGATGAATAAAGCTCCCTGGGGGGACCCGGAC	960
Db	1038	CAGAAGCTTCCCACAGAAGCTCCAACGGCTGATGAATAAAGCTCCCTGGGGGGACCCGGAC	1097
QY	961	TGCTCCCCGGACCAATAGGGGGGTGACAGAGTCTGGCTAG	1002
Db	1098	TGCTCCCCGGACCAATAGGGGGGTGACAGAGTCTGGCTAG	1130

DECEMBER 1999

RESOLUTION
US-09-764-847-1782

Sequence 1782, Application US/09764847

Patent NO. US20020132767A1

APPLICANT: Rosen et al.

FILE OF INVENTION: Nucleic Acids, Pr
FILE REFERENCE: PC009

CURRENT APPLICATION NUMBER: US/09/764

Prior application data removed - cons

NUMBER OF SEQ ID NOS: 2003
SOFTWARE: Datascan 2.0

SOFTWARE: FALCONCHM VER. 2.0
SEQ ID NO 1782

LENGTH: 2651
TYPE: DNA

ORGANISM: Homo sapiens

JS-09-764-847-1782

Query Match	Score
57.6%	

Best Local Similarity 100.0%; Pred.
Matches 577; Conserved 0

```

CONSERVATIVE; M18

```

1 ATGTGAGCTGCAGCTGGTTCACCGC

2075 ATGTGAGCTGCAGCTGCTCAACGGC

61 GARNETT, G. L. 1963. The ecology of the

[illegible]

2135 GACCTGCAGCTGGGCTGTCACTGTTG

Y 121 GGCCTGTGCTACACGCCCTGCTGGTG

b
2195 GGCCTGTTACAAAGCCATTCACCATTCA

[illegible]

QY 181 ATGCCGAGCGTGTACTTTTGTCAACATGAGCAGGAGCCGTTGCTCAGGCGCCCTGACC 240
DB 2255 ATGCCGAGCGTGTACTTTTGTCAACATGAGCAGGAGCCGTTGCTCAGGCGCCCTGACC 2314
QY 241 CCTGTGACCTGTCTGAGCCCGCCGAGCTCCCGTGTGGCGCTGTGAGTGTGGCGGAA 300
DB 2315 CCTGTGACCTGTCTGAGCCCGCCGAGCTCCCGTGTGGCGCTGTGAGTGTGGCGGAA 2374
QY 301 GTCCACGTGGCAGTGCAGATCCCTTCAATGTCTCTCACTGTGGCCATGTACTCAAC 360
DB 2375 GTCCACGTGGCAGTGCAGATCCCTTCAATGTCTCTCACTGTGGCCATGTACTCAAC 2434
QY 361 GCCCTGTGAGCCCTCGACCACTACATGAGGCTGACCTGCGGACCTACATGCGCAGC 420
DB 2435 GCCCTGTGAGCCCTCGACCACTACATGAGGCTGACCTGCGGACCTACATGCGCAGC 2494
QY 421 GTGTACAACACGCGGACAGTGTGGCGCTTCTGTGGGGTGGCGGCTGTGACAGCTTC 480
DB 2495 GTGTACAACACGCGGACAGTGTGGCGCTTCTGTGGGGTGGCGGCTGTGACAGCTTC 2554
QY 481 TCCTCGCTGCTCTTCTACATCTGCAAGCCATGTCTCACCCGCGCTAGAGTGGCCAG 540
DB 2555 TCCTCGCTGCTCTTCTACATCTGCAAGCCATGTCTCACCCGCGCTAGAGTGGCCAG 2614
QY 541 ATGAGAACGAGAAAGCTGCGAGCCAGCGCTGTGT 577
DB 2615 ATGAGAACGAGAAAGCTGCGAGCCAGCGCTGTGT 2651

RESULT 14

US-10-092-154-1782
; Sequence 1782, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC093C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1782
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1782

Query Match 57.6%; Score 577; DB 14; Length 2651;
Best Local Similarity 100.0%; Pred. No. 1.6e-146; Mismatches 0; Gaps 0;

Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTGAGCTGCAAGCTGTTCAACGAGCAGGAGTGTGAGAGAGTGTGCTGCTGACG 60
DB 2075 ATGTGAGCTGCAAGCTGTTCAACGAGCAGGAGTGTGAGAGAGTGTGCTGCTGACG 2134
QY 61 GACCTGAGCTGAGGCTGTCACTGTGTGCTGTGAGGCTGTGTGTGTGTGTGTGTGTGT 120
DB 2135 GACCTGAGCTGAGGCTGTCACTGTGTGCTGTGAGGCTGTGTGTGTGTGTGTGTGTGT 2194
QY 121 GGGCTGTGCTACAAAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
DB 2195 GGGCTGTGCTACAAAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2254
QY 181 ATGCCGAGCGTGTACTTTTGTCAACATGAGCAGGAGCCGTTGCTCAGGCGCCCTGACC 240
DB 2255 ATGCCGAGCGTGTACTTTTGTCAACATGAGCAGGAGCCGTTGCTCAGGCGCCCTGACC 2314
QY 241 CCTGTGACCTGTCTGAGCCCGCCGAGCTCCCGTGTGGCGCTGTGAGTGTGGCGGAA 300
DB 2315 CCTGTGACCTGTCTGAGCCCGCCGAGCTCCCGTGTGGCGCTGTGAGTGTGGCGGAA 2374

QY 301 GTCCACGTGGCAGTGCAGATCCCTTCAATGTCTCTCACTGTGGCCATGTACTCAAC 360
DB 2375 GTCCACGTGGCAGTGCAGATCCCTTCAATGTCTCTCACTGTGGCCATGTACTCAAC 2434
QY 361 GCCCTGTGAGCCCTCGACCACTACATGAGGCTGACCTGCGGACCTACATGCGCAGC 420
DB 2435 GCCCTGTGAGCCCTCGACCACTACATGAGGCTGACCTGCGGACCTACATGCGCAGC 2494
QY 421 GTGTACAACACGCGGACAGTGTGGCGCTTCTGTGGGGTGGCGGCTGTGACAGCTTC 480
DB 2495 GTGTACAACACGCGGACAGTGTGGCGCTTCTGTGGGGTGGCGGCTGTGACAGCTTC 2554
QY 481 TCCTCGCTGCTCTTCTACATCTGCAAGCCATGTCTCACCCGCGCTAGAGTGGCCAG 540
DB 2555 TCCTCGCTGCTCTTCTACATCTGCAAGCCATGTCTCACCCGCGCTAGAGTGGCCAG 2614
QY 541 ATGAGAACGAGAAAGCTGCGAGCCAGCGCTGTGT 577
DB 2615 ATGAGAACGAGAAAGCTGCGAGCCAGCGCTGTGT 2651

RESULT 15

US-09-918-995-8117
; Sequence 8117, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8117
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-8117

Query Match 35.8%; Score 358.6; DB 10; Length 398;
Best Local Similarity 94.0%; Pred. No. 2.3e-87; Mismatches 24; Indels 0; Gaps 0;

Matches 373; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 380 ACTACATGAGCGTGTGACCTGCGGACCTACATGTGCGCAGCGTGTACAAACGCGGACG 439
DB 1 ACTACATGAGCGTGTGACCTGCGGACCTACATGTGCGCAGCGTGTACAAACGCGGACG 60
QY 440 TGTGCGGCTTGT 499
DB 61 TGTGCGGCTTGT 120
QY 500 TCTGAGCCATGTGTCCACCGGCGGCTGTAGAGCGCAAGATGAGAGAGAGAGAGAGAG 559
DB 121 TCTGAGCCATGTGTCCACCGGCGGCTGTAGAGCGCAAGATGAGAGAGAGAGAGAGAG 180
QY 560 CCGAGCGCAGCTGT 619
DB 181 CCGAGCGCAGCTGT 240
QY 620 TGTGTCTACTTCTCCCGGCTCGAGAGAGAGACAGCCCTGTGACCGGAGACAGGCGCG 679
DB 241 TGTGTCTACTTCTCCCGGCTCGAGAGAGAGACAGCCCTGTGACCGGAGACAGGCGCG 300
QY 680 TGGAGCCCTGAGGACAGGCTGT 739
DB 301 TGGAGCCCTGAGGACAGGCTGT 360
QY 740 CGCCACACTATCTGATCTGT 776
DB 361 CGCCACACTATCTGATCTGT 397

Wed Dec 15 11:04:29 2004

us-09-995-225b-1.rmpb

Page 12

Search completed: December 15, 2004, 02:53:08
Job time : 514.891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:17:59 ; Search time 3155.64 Seconds
(without alignments)
11570.599 Million cell updates/sec

Title: US-09-995-225B-1

Perfect score: 1002
Sequence: 1 atctggagctgcagctgcgtc.....tcgacgagctgcgtgcgtag 1002

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_nc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g881.*
9: gb_g882.*

Pred. No. is the number of results predicted to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1002	100.0	1002	9 AY403440	AY403440 Homo sapi
2	1002	100.0	1787	3 CR605880	CR605880 full-length
3	1002	100.0	1809	3 CR619529	CR619529 full-length
4	1002	100.0	1861	3 CR619773	CR619773 full-length
5	883.8	88.2	1094	5 BX422717	BX422717 BX422717
6	843.6	83.8	1042	5 BX448207	BX448207 BX448207
7	839.2	83.8	1084	5 BX463176	BX463176 BX463176
8	757	75.5	1001	5 BX338900	BX338900 BX338900
9	742.8	74.1	1023	5 BX382511	BX382511 BX382511
10	665	66.4	907	5 BX332385	BX332385 BX332385
11	653.8	65.2	670	9 AY403441	AY403441 Pan trogl
12	647	64.6	851	4 B1825260	B1825260 603071957
13	645.4	64.4	996	5 BQ712494	BQ712494 AGENCOURT
14	642	64.1	785	4 B1818819	B1818819 603037676
15	610	60.9	996	9 AY403442	AY403442 Mus muscu
16	610	60.9	1737	3 AKO89232	AKO89232 Mus muscu
17	610	60.9	1758	3 AKO46512	AKO46512 Mus muscu
18	608.4	60.7	1428	4 AKO87100	AKO87100 Mus muscu
19	598	59.7	773	4 B1910046	B1910046 603068162
20	596	59.5	1429	3 AKO50034	AKO50034 Mus muscu
21	592.2	59.1	925	4 BG326599	BG326599 602425455
22	586.2	58.5	940	4 B1754831	B1754831 603025292
23	581	58.0	709	4 B1820327	B1820327 603036501
24	549.8	54.9	826	4 B1818174	B1818174 603032623

25	542.4	54.1	544	4 BG570221	BG570221 602590622
26	539.8	53.9	650	2 BE731686	BE731686 601567082
27	512.6	51.2	914	5 BX377255	BX377255 BX377255
28	502	50.1	609	6 CD676083	CD676083 EB3108.Y
29	501.6	50.1	824	4 B1767814	B1767814 603060921
30	467.6	46.7	699	4 BM007558	BM007558 603616944
31	443	44.3	1165	4 BM547338	BM547338 AGENCOURT
32	431.4	43.1	733	4 B1489911	B1489911 603031665
33	398.6	39.8	773	6 CB990075	CB990075 AGENCOURT
34	393.6	39.3	703	4 B1819918	B1819918 603035463
35	390	38.9	635	7 CF738689	CF738689 UI-M-HD0-
36	376.4	37.6	657	8 AZ801707	AZ801707 2M0060G16
37	372.4	37.2	697	7 CK836653	CK836653 4061701.B
38	364.6	36.4	657	6 BY747823	BY747823 BY747823
39	364	36.3	597	1 A1189791	A1189791 ME31C04.T
40	364	36.3	665	2 BB653568	BB653568 BB653568
41	360	35.9	499	5 BX488779	BX488779 DKF2P686O
42	353	35.2	692	4 BJ735246	BJ735246 BJ735246
43	352.8	35.2	886	7 CR587202	CR587202 CR587202
44	351.2	35.0	629	2 BB664126	BB664126 BB664126
45	349.4	34.9	808	5 BP711583	BP711583 BP711583

ALIGNMENTS

RESULT 1
AY403440
LOCUS
DEFINITION Homo sapiens HCM1558 gene, VIRUAL TRANSCRIPT, partial sequence,
ACCESSION AY403440
VERSION AY403440.1 GI:39759423
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1002)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Clevello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smutsky,J.J.,
Adams,M.D. and Cargill,M.
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene titlos
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Clevello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smutsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. 1002
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1002
/locus_tag="HCM1558"
ORIGIN
Query Match 100.0%; Pred. No. 1002; DB 9; Length 1002;
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 ATGTGAGCTGCACTGCTTCAACGACAGGCTGTGAGAGAGCTGCTGCTCCAG 60
Db 1 ATGTGAGCTGCACTGCTTCAACGACAGGCTGTGAGAGAGCTGCTGCTCCAG 60

QY 61 GACCTGACAGTGGGGGCTGTCACTGTTTGTCTGCTGGGCTGTGGTGGGCTGGCAATG 120
 Db 61 GACCTGACAGTGGGGGCTGTCACTGTTTGTCTGCTGGGCTGTGGTGGGCTGGCAATG 120
 QY 121 GGCTGTGTGTACACGCGCTGTGGTGTGGCCACTACACAGCAAGGCCAGCATGACC 180
 Db 121 GGCTGTGTGTACACGCGCTGTGGTGTGGCCACTACACAGCAAGGCCAGCATGACC 180
 QY 181 ATGCCGAGGTGTACTTTTGTCAACTGGCAGTGGCAGGCTGTGGTCTCAGGCGCCCTGACC 240
 Db 181 ATGCCGAGGTGTACTTTTGTCAACTGGCAGTGGCAGGCTGTGGTCTCAGGCGCCCTGACC 240
 QY 241 CCTGTGACCTGTGGGCGCCCGGAGCTCCCGGTGGGGGCTGTGGAGTGGTGGGGCGAA 300
 Db 241 CCTGTGACCTGTGGGCGCCCGGAGCTCCCGGTGGGGGCTGTGGAGTGGTGGGGCGAA 300
 QY 301 GTTCCACGTGGGCACTGCAAGATCCCTTCAATGTCTTCACTGTGGCCATGTACTCCACC 360
 Db 301 GTTCCACGTGGGCACTGCAAGATCCCTTCAATGTCTTCACTGTGGCCATGTACTCCACC 360
 QY 361 GCGCTGTGAGCTGTGACCACTAATGAGGTGCACTGCGCGGGAGCTTACATGAGCCAGC 420
 Db 361 GCGCTGTGAGCTGTGACCACTAATGAGGTGCACTGCGCGGGAGCTTACATGAGCCAGC 420
 QY 421 GTGTACACACGCGGCACTGTGTGGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 480
 Db 421 GTGTACACACGCGGCACTGTGTGGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 480
 QY 481 TCCTGTGCTCTTCTTCACTGTGAGGCACTGTGTGCACTGCGCGGCTGTGAGTGGCCAG 540
 Db 481 TCCTGTGCTCTTCTTCACTGTGAGGCACTGTGTGCACTGCGCGGCTGTGAGTGGCCAG 540
 QY 541 ATGCAGAACGAGAAAGCTGCGGACGCGCACTGCTGCTTCACTGCGGCTTCACTGCTGAG 600
 Db 541 ATGCAGAACGAGAAAGCTGCGGACGCGCACTGCTGCTTCACTGCGGCTTCACTGCTGAG 600
 QY 601 CTGGCCACCTCTTACGCGCTGTGTGTCTTCCCGCTGTGGAGGAGGACAGCCCTG 660
 Db 601 CTGGCCACCTCTTACGCGCTGTGTGTCTTCCCGCTGTGGAGGAGGACAGCCCTG 660
 QY 661 GACCGGAGACAGGCGCGGCTGTGAGGCGCTGTGGGACACAGGCTGTGTGCGACCTGTG 720
 Db 661 GACCGGAGACAGGCGCGGCTGTGAGGCGCTGTGGGACACAGGCTGTGTGCGACCTGTG 720
 QY 721 ACCGATTTGGGCTGTGAGCGGCACTATCTGATCTGTGCTGTGGGACACGCTCATATC 780
 Db 721 ACCGATTTGGGCTGTGAGCGGCACTATCTGATCTGTGCTGTGGGACACGCTCATATC 780
 QY 781 TCGCAGAGGAAAGCCGCTGTGAGCGCACTACCTGGGCTTCTGCACTTTGTGAAGATTC 840
 Db 781 TCGCAGAGGAAAGCCGCTGTGAGCGCACTACCTGGGCTTCTGCACTTTGTGAAGATTC 840
 QY 841 TCCAACTCTGTGCTTCTTCCAGCAGCTTGTGACACCACTTCTTACCGCTACATGAG 900
 Db 841 TCCAACTCTGTGCTTCTTCCAGCAGCTTGTGACACCACTTCTTACCGCTACATGAG 900
 QY 901 CAGAGTTTCCCGAGAAAGCTTCCAGCGCTGATGAAAGCTGCTGTGGGAGCCGAGC 960
 Db 901 CAGAGTTTCCCGAGAAAGCTTCCAGCGCTGATGAAAGCTGCTGTGGGAGCCGAGC 960
 QY 961 TGTCTCCCGGACCAATGTGGGGTGTGACGAGGTGTGGGGTAG 1002
 Db 961 TGTCTCCCGGACCAATGTGGGGTGTGACGAGGTGTGGGGTAG 1002

RESULT 2
 CR605880 1787 bp mRNA linear HTC 21-JUL-2004
 LOCUS
 DEFINITION full-length cDNA clone CS0DN001Yc15 of Adult brain of Homo sapiens
 ACCESSION CR605880
 VERSION CR605880.1 GI:50486687

KEYWORDS
 SOURCE
 ORGANISM
 HOMO SAPIENS (human)
 HOMO SAPIENS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 (bases 1 to 1787)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Feng Liang Email: fliang@life.technet.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue
 REFERENCE
 2 (bases 1 to 1787)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
 Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 LOCATION/Qualifiers
 1..1787
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DN001Yc15"
 /tissue_type="Adult brain"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Query Match 100.0%; Score 1002; DB 3; Length 1787;
 Best Local Similarity 100.0%; Pred. No. 1, 2e-213;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTGAGCTTCACCTGCTTCAACGCGAGGCTGTGGAGAGCTGCTGCTGCGCAG 60
 Db 71 ATGTGAGCTTCACCTGCTTCAACGCGAGGCTGTGGAGAGCTGCTGCTGCGCAG 130
 QY 61 GACCTGAGCTGGGGGCTGTGACCTGTTGTGCTGTGGGCTGTGGGCTGTGGGCTGTGG 120
 Db 131 GACCTGAGCTGGGGGCTGTGACCTGTTGTGCTGTGGGCTGTGGGCTGTGGGCTGTGG 190
 QY 121 GGCTGTGTGTACACGCGCTGTGGTGTGGCCACTACACAGCAAGGCCAGCATGACC 180
 Db 191 GGCTGTGTGTACACGCGCTGTGGTGTGGCCACTACACAGCAAGGCCAGCATGACC 250
 QY 181 ATGCCGAGGTGTACTTTTGTCAACTGGCAGTGGCAGGCTGTGGTCTCAGGCGCCCTGACC 240
 Db 251 ATGCCGAGGTGTACTTTTGTCAACTGGCAGTGGCAGGCTGTGGTCTCAGGCGCCCTGACC 310
 QY 241 CCTGTGACCTGTGGGCGCCCGGAGCTCCCGGTGGGGGCTGTGGAGTGGGGCGAA 300
 Db 311 CCTGTGACCTGTGGGCGCCCGGAGCTCCCGGTGGGGGCTGTGGAGTGGGGCGAA 370
 QY 301 GTTCCACGTGGGCACTGCAAGATCCCTTCAATGTCTTCACTGTGGCCATGTACTCCACC 360
 Db 371 GTTCCACGTGGGCACTGCAAGATCCCTTCAATGTCTTCACTGTGGCCATGTACTCCACC 430
 QY 361 GCGCTGTGAGCTGTGACCACTAATGAGGTGCACTGCGCGGGAGCTTACATGAGCCAGC 420
 Db 431 GCGCTGTGAGCTGTGACCACTAATGAGGTGCACTGCGCGGGAGCTTACATGAGCCAGC 490
 QY 421 GTGTACACACGCGGCACTGTGTGGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 480
 Db 491 GTGTACACACGCGGCACTGTGTGGGCTTGTGTGGGGTGGCGGCTGTGACCAAGCTTC 550
 QY 481 TCCTGTGCTCTTCTTCACTGTGAGGCACTGTGTGCACTGCGCGGCTGTGAGTGGCCAG 540
 Db 551 TCCTGTGCTCTTCTTCACTGTGAGGCACTGTGTGCACTGCGCGGCTGTGAGTGGCCAG 610

Qy	541	ATGCAAGAACGCAAGAGCTGCGCAGCACAAGCTGGGTTCATACGATACGTGGTGACAGA	600
Db	611	ATGCAAGAACGCAAGAGCTGCGCAGCACAAGCTGGGTTCATACGATACGTGGTGACAGA	670
Qy	601	CTGGCCACCTCTACGCGCTGGTGTACTCTTCCGCGTCCGACAGGAGGACACGCCCTCG	660
Db	671	CTGGCCACCTCTACGCGCTGGTGTACTCTTCCGCGTCCGACAGGAGGACACGCCCTCG	730
Qy	661	GACCGGGGACAGGGGCGCGGTGGAGCCCTCCGACACAAGGCTGCTGGTGGCCACCGTGGC	720
Db	731	GACCGGGGACAGGGGCGCGGTGGAGCCCTCCGACACAAGGCTGCTGGTGGCCACCGTGGC	790
Qy	721	ACGCAGTTTGGGCTCTGGACGCAACACTATCTGATCTGTGGGCAACGGATCATCTC	780
Db	791	ACGCAGTTTGGGCTCTGGACGCAACACTATCTGATCTGTGGGCAACGGATCATCTC	850
Qy	781	TCCGGAGGGAAAGCCCGTGGAGGCAACACTACTCTGGGGGCTACTGCACTTTGTGAAGATTTC	840
Db	851	TCCGGAGGGAAAGCCCGTGGAGGCAACACTACTCTGGGGGCTACTGCACTTTGTGAAGATTTC	910
Qy	841	TCCAAATCTCTGGGCTTCTCCAGCAGCTTGTGACACACCACTTCTCTACCGCTACATGAAC	900
Db	911	TCCAAATCTCTGGGCTTCTCCAGCAGCTTGTGACACACCACTTCTCTACCGCTACATGAAC	970
Qy	901	CAGAGCTTCCCGACAGCTTCCAAACGCGTGAATGAAGAAAGTGCCTCTGCGGGACCCGAC	960
Db	971	CAGAGCTTCCCGACAGCTTCCAAACGCGTGAATGAAGAAAGTGCCTCTGCGGGACCCGAC	1030
Qy	961	TGCTTCCCGGACACATGGGGGTGAGAGAGTGCCTGGGGTGA	1002
Db	1031	TGCTTCCCGGACACATGGGGGTGAGAGAGTGCCTGGGGTGA	1072

RESULT 3	LOCUS	DEFINITION	1809 bp	mRNA	linear	HTC 21-JUL-2004
CR619529	CR619529	full-length cDNA clone CS0DM006YF08 of Fetal liver of Homo sapiens (human).				
ACCESSION	CR619529					
VERSION	CR619529.1	GI:50500336				
KEYWORDS	HTC; CNSLT_cDNA.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 1809)					
AUTHORS	Ll,W.B., Gruber,C., Jessee,J. and Polayes,D.					
TITLE	Full-length cDNA libraries and normalization					
JOURNAL	Unpublished					
REMARK	Contact : Feng Liang Email : fliang@life tech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue 2 (bases 1 to 1809)					
REFERENCE	Genoscope.					
AUTHORS	Direct Submission					
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :					
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr					
COMMENT	- Web : www.genoscope.cns.fr					
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers					
	end enriched, double-strand cDNA was digested with Not I and cloned					
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library					
	was normalized. Library was constructed by Life Technologies, a					
	division of Invitrogen.					

FEATURES	Location/Qualifiers
Source	1. .1809

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM006P08"
/tissue_type="Fetal liver"
/plasmid="pCMVSPORT_6"

```

ORIGIN

Query Match	100.0%;	Score 1002;	DB 3;	Length 1809;
Best Local Similarity	100.0%;	Pred. No. 1.2e-213;		
Matches 1002; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]

RESULT 4
CR619773
LOCUS
DEFINITION full-length cDNA clone CS0DM007YN21 of Fetal liver of Homo sapiens (human).
ACCESSION CR619773
VERSION CR619773.1 GI:50500580
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1861)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/InvitrogenCorporation1600
REFERENCE 2 (bases 1 to 1861)
AUTHORS Faraday Avenue
TITLE Genoscope.
JOURNAL Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
Source Location/Qualifiers
1..1861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007YN21"
/issue_type="Fetal liver"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 100.0%; Score 1002; DB 3; Length 1861;
Best Local Similarity 100.0%; Pred. No. 1.3e-213;
Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGAGCTGACGCTGTTTCAACGACAGAGGCTGTGAGAGAGCTGCTGCTCCAG 60
DB 124 ATGTGAGCTGACGCTGTTTCAACGACAGAGGCTGTGAGAGAGCTGCTGCTCCAG 183
QY 61 GACCTGACGCTGGGGCTGTCACCTGTTGTGCTGCTGGGCTGTGCTGGCTGCCAGT 120
DB 184 GACCTGACGCTGGGGCTGTCACCTGTTGTGCTGCTGGGCTGTGCTGGCTGCCAGT 243
QY 121 GGGCTGTGTCACACCGCTGTGCTGTGCTGCAACCTTACACAGCAAGGCGACATGACC 180
DB 244 GGGCTGTGTCACACCGCTGTGCTGTGCTGCGCAACTTACACAGCAAGGCGACATGACC 303
QY 181 ATGCCGAGAGTGTACTTTGTCAACATGAGAGTGGCAAGGCTGTGCTCAACGCGCTTGCC 240
DB 304 ATGCCGAGAGTGTACTTTGTCAACATGAGAGTGGCAAGGCTGTGCTCAACGCGCTTGCC 363
QY 241 CCTGTGACCTGCTCGGCGCCCGGAGCTCCGGTGGGGCTGTGAGAGTGGGCGCGCA 300
DB 364 CCTGTGACCTGCTCGGCGCCCGGAGCTCCGGTGGGGCTGTGAGAGTGGGCGCGCA 423
QY 301 GTTCAAGTGGCACTGTCAGATCCCTTCAATGTGTCTTCACTGTGAGGCACTGATCCACC 360
DB 424 GTTCAAGTGGCACTGTCAGATCCCTTCAATGTGTCTTCACTGTGAGGCACTGATCCACC 483
QY 361 GCGCGCTGAGGCTGACCACTACATGAGGCTGACCTGCGCGGACCTTACATGGCCAGC 420
DB 484 GCGCGCTGAGGCTGACCACTACATGAGGCTGACCTGCGCGGACCTTACATGGCCAGC 543
QY 421 GTTCAACACGCGGCACTGTGCGGCTTCTGTGGGGTGGCGGCTGTACCAACGCTTC 480

DB 544 GTTCAACACGCGGCACTGTGCGGCTTCTGTGGGGTGGCGGCTGTACCAACGCTTC 603
QY 481 TCCTCGCTGCTCTTCAACATCTGACAGCAGATGTCCACCGGCGCTTACATGGCCAG 540
DB 604 TCCTCGCTGCTCTTCAACATCTGACAGCAGATGTCCACCGGCGCTTACATGGCCAG 663
QY 541 ATGCAAGACGACAGAGCTCCGACGCGCACTGTGTTTCATGAGCTACGCTGTCACGA 600
DB 664 ATGCAAGACGACAGAGCTCCGACGCGCACTGTGTTTCATGAGCTACGCTGTCACGA 723
QY 601 CTGGGCAACCTTACAGCGCTGTGCTACTCTCCGCTCCGACGAGAGACACGCGCTTG 660
DB 724 CTGGGCAACCTTACAGCGCTGTGCTACTCTCCGCTCCGACGAGAGACACGCGCTTG 783
QY 661 GACCGGACACGCGGCGGCTGTGAGCGCTTGTGACACAGGCTGTGTTGGCCACGCTGTC 720
DB 784 GACCGGACACGCGGCGGCTGTGAGCGCTTGTGACACAGGCTGTGTTGGCCACGCTGTC 843
QY 721 AGCAGTTTGGGCTTGTGACGCGCACATCTGATCTGCTGGGACACGCTCATCATC 780
DB 844 AGCAGTTTGGGCTTGTGACGCGCACATCTGATCTGCTGGGACACGCTCATCATC 903
QY 781 TCGCGAGGAGACCGGCTGTGACGCGCACATCTGAGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 904 TCGCGAGGAGACCGGCTGTGACGCGCACATCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 963
QY 841 TCCAAACTCTGCTGCTTGTGACGAGCTTGTGACACCACTTCTTCAACGCTGCTGCTGCT 900
DB 964 TCCAAACTCTGCTGCTTGTGACGAGCTTGTGACACCACTTCTTCAACGCTGCTGCTGCT 1023
QY 901 CAGACCTTCCCGACGACACTTCAACGCTGATGAAAAAGCTGCTGCGGAGACCGGAC 960
DB 1024 CAGACCTTCCCGACGACACTTCAACGCTGATGAAAAAGCTGCTGCGGAGACCGGAC 1083
QY 961 TGTCTCCCGGACCAATGAGGCTGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1002
DB 1084 TGTCTCCCGGACCAATGAGGCTGACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
RESULT 5
BX422717
LOCUS BX422717
DEFINITION BX422717 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone CS0DM001YX15 5-PRIME, mRNA sequence.
ACCESSION BX422717
VERSION BX422717.2 GI:46927958
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1094)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30643222.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 4338.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DM001AB08P1c=4338.f.
FEATURES
Source Location/Qualifiers
1..1094
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="CS0DN001YC15"
/issue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 88.2%; Score 883.8; DB 5; Length 1094;
Best Local Similarity 95.4%; Pred. No. 3.4e-187;
Matches 949; Conservative 13; Mismatches 25; Indels 8; Gaps 5;

QY 1 ATGTGAGCTGACGCTGTTTCAACGACGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 60
DB ATGTGAGCTGACGCTGTTTCAACGACGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 129

QY 61 GACCTGACGCTGGGGCTGTCACTGTTTCCGCTGCTGCGCTGTGTGGGCTGTGCCAGT 120
DB GACCTGACGCTGGGGCTGTCACTGTTTCCGCTGCTGCGCTGTGTGGGCTGTGCCAGT 189

QY 121 GGCCTGTGCTACCAACGCGCTGTGTGCTGAGCCACTACACAGCAAGGCGCAGATGACC 180
DB GGCCTGTGCTACCAACGCGCTGTGTGCTGAGCCACTACACAGCAAGGCGCAGATGACC 249

QY 181 ATGCCGAGCTGTATCTTTGTCAACATGAGCAGTGTGAGCAGCTGTGCTCAGCGCCTGACC 240
DB ATGCCGAGCTGTATCTTTGTCAACATGAGCAGTGTGAGCAGCTGTGCTCAGCGCCTGACC 309

QY 241 CCTGTGACCTGCTGCGCCCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAGAA 300
DB CCTGTGACCTGCTGCGCCCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGAGAA 369

QY 301 GTCCAGTGTGACCTGACGATCCCGCTTCAATGATGTCTCATCGGAGGCGCATGTACTCCACC 360
DB GTCCAGTGTGACCTGACGATCCCGCTTCAATGATGTCTCATCGGAGGCGCATGTACTCCACC 429

QY 361 GGCCTGTGAGCTGTGACCACTATGAGCAGTGTGAGCAGCTGTGCGGAGCACTATGAGCCAGC 420
DB GGCCTGTGAGCTGTGACCACTATGAGCAGTGTGAGCAGCTGTGCGGAGCACTATGAGCCAGC 489

QY 421 GTGTACCAACGCGGCGCAGTGTGCGCTTGTGTGGGCTGTGGCGCTGTGACCAAGTTTC 480
DB GTGTACCAACGCGGCGCAGTGTGCGCTTGTGTGGGCTGTGGCGCTGTGACCAAGTTTC 549

QY 481 TCCTGCTGCTCTTCAATCTGACATGTGTGACATGTGTGACATGTGTGACATGTGTGACATGT 540
DB TCCTGCTGCTCTTCAATCTGACATGTGTGACATGTGTGACATGTGTGACATGTGTGACATGT 609

QY 541 ATGCAGAACGACAGAGCTGTGAGCAGCAGCAGCTGTGTTCATCGGCTACGTTGTCAGCA 600
DB ATGCAGAACGACAGAGCTGTGAGCAGCAGCAGCTGTGTTCATCGGCTACGTTGTCAGCA 669

QY 601 CTGGCCACCTCTTACGCGCTGTGTGCTACTTCCCGCTGCGCAGGGA-6GACAGCCCTT 659
DB CTGGCCACCTCTTACGCGCTGTGTGCTACTTCCCGCTGCGCAGGGA-6GACAGCCCTT 729

QY 660 GAGCGCGGAGACGCGGCGCTGTGAGCAGCAGCAGCTGTGTGAGCAGCAGCAGCTGTGTG 719
DB GAGCGCGGAGACGCGGCGCTGTGAGCAGCAGCAGCTGTGTGAGCAGCAGCAGCTGTGTG 789

QY 720 CACGCAAGTTT-6GCGCTGTGAGCAGCAGCAGCTATCTGCTGTGAGGAGCAGCAGCTATCA 778
DB CACGCAAGTTT-6GCGCTGTGAGCAGCAGCAGCTATCTGCTGTGAGGAGCAGCAGCTATCA 849

QY 779 TCTGCGAGAGGAGAGCCGCTGTGAGCAGCAGCTATCTGCGGCTGTGAGCAGCTTGTGAGAGATT 838
DB TCTGCGAGAGGAGAGCCGCTGTGAGCAGCAGCTATCTGCGGCTGTGAGCAGCTTGTGAGAGATT 909

QY 839 TCTCCAAACCTCTGCGCTTCTCAGAGCTTGTGACACACACTTCTTACCGGCTACATGA 898

DB 910 TCTCCAAACCTCTGCGCTTCTCAGAGCTTGTGACACACACTTCTTACCGGCTACATGA 969

QY 899 ACCAGAGCTTCCCGCAGACGCTTCAACGAGCTATGAAAAGCTCCCTGCGGAGACCGGC 958

DB 970 AC--AAGAGCTTCCCGCAGACGCTTCAACGAGCTATGAAAAGCTCCCTGCGGAGACCGGC 1023

QY 959 ACTGCTCCCGCAGACGCTTCAACGAGCTATGAAAAGCTCCCTGCGGAGACCGGC 993

DB 1024 TCTCCAAACCTCTGCGCTTCTCAGAGCTTGTGACACACACTTCTTACCGGCTACATGA 1058

RESULT 6
BX448207 1042 bp mRNA linear EST 06-MAY-2004
LOCUS
DEFINITION
BX448207 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM006YP08 5-PRIME, mRNA sequence.
ACCESSION
BX448207
VERSION
BX448207.2 GI:47071231
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
Li, W.B., Gruber, C., Jesse, J. and Polyes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31033927.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4338.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DM006H040P1&c=4338.f.

FEATURES

source
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM006YP08"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 84.2%; Score 843.6; DB 5; Length 1042;
Best Local Similarity 97.0%; Pred. No. 3.3e-178;
Matches 883; Conservative 7; Mismatches 17; Indels 3; Gaps 3;

QY 1 ATGTGAGCTGACGCTGTTTCAACGACGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 60
DB ATGTGAGCTGACGCTGTTTCAACGACGAGGCTGTGAGAGAGCTGCTGCTGCGCAG 134

QY 61 GACCTGACGCTGGGGCTGTCACTGTTTCCGCTGCTGCGCTGTGTGGGCTGTGCCAGT 120
DB GACCTGACGCTGGGGCTGTCACTGTTTCCGCTGCTGCGCTGTGTGGGCTGTGCCAGT 194

QY 121 GGCCTGTGCTACCAACGCGCTGTGTGCTGAGCCACTACACAGCAAGGCGCAGATGACC 180
DB GGCCTGTGCTACCAACGCGCTGTGTGCTGAGCCACTACACAGCAAGGCGCAGATGACC 254

```
OY 181 ATGCCGAGCTGATCTTTGTCAACATGGCAGTGGCGCTGTGTCTGACAGGCCCTTGACC 240
DB 255 ATGCCGAGCTGATCTTTGTCAACATGGCAGTGGCGCTGTGTCTGACAGGCCCTTGACC 314
OY 241 CCTGTGACCTGTGTGGCCCCCGGAGCTCCCGATGGGCGCTGTGGAGTGTGGCGCGCA 300
DB 315 CTTGTGACCTGTGTGGCCCCCGGAGCTCCCGATGGGCGCTGTGGAGTGTGGCGCGCA 374
OY 301 GTCCAGTGTGACCTGTGGAGTCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCCACC 360
DB 375 GTCCAGTGTGACCTGTGGAGTCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCCACC 434
OY 361 GCCCTGTGAGCTGTGACCACTACATGAGCGGCACTGCGCGGCACTACATGGCCAGC 420
DB 435 GCCCTGTGAGCTGTGACCACTACATGAGCGGCACTGCGCGGCACTACATGGCCAGC 494
OY 421 GTGTACAACACGCGGCACTGTGTGCGGCTTGTGTGGGATGTGGCGGCTGTGACCACTTC 480
DB 495 GTGTACAACACGCGGCACTGTGTGCGGCTTGTGTGGGATGTGGCGGCTGTGACCACTTC 554
OY 481 TCCTGTGCTCTTTTACATCTGTGACGATGTGTCCACCGCGGCGCTAGAGTGGCGCAAG 540
DB 555 TCCTGTGCTCTTTTACATCTGTGACGATGTGTCCACCGCGGCGCTAGAGTGGCGCAAG 614
OY 541 ATGAGAGACGAGAGGCTGCGGCACTGCGGCTGTGTTCATCGGCTACGTTGTGCGCAGCA 600
DB 615 ATGAGAGACGAGAGGCTGCGGCACTGCGGCTGTGTTCATCGGCTACGTTGTGCGCAGCA 674
OY 601 CTGGCCACCTCTTACGCGCTGTGTCTACTCTCCCGCTTCCGACGAGAGACACGCCCCC 660
DB 675 CTGGCCACCTCTTACGCGCTGTGTCTACTCTCCCGCTTCCGACGAGAGACACGCCCCC 734
OY 661 GACCGGACACAGGCGCGGCTGTGAGCCCTCGGCAACAGGCTGTGTGGCGCACTGTTGC 720
DB 735 GACCGGACACAGGCGCGGCTGTGAGCCCTCGGCAACAGGCTGTGTGGCGCACTGTTGC 794
OY 721 ACGCAGTTGGGCTCTGTGACGCGCACTATCTGATCTCTGTGGGCAACGCTCATATC 780
DB 795 ACGCAGTTGGGCTCTGTGACGCGCACTATCTGATCTCTGTGGGCAACGCTCATATC 853
OY 781 TCGCAGAGGAGACCGCTGTGAGCACTACCTGTGGGCTACTGCACTTTGTGAAGATTC 840
DB 854 TCGCAGAGGAGACCGCTGTGAGCACTACCTGTGGGCTACTGCACTTTGTGAAGATTC 913
OY 841 TCCTAATCTCTGCTTCTCCAGCAGCTTTGTGACCACTTCTCTACCGCTACATGAC 900
DB 914 TCCTAATCTCTGCTTCTCCAGCAGCTTTGTGACCACTTCTCTACCGCTACATGAC 972
OY 901 CAGAGCTTCC 910
DB 973 AGAAGCTTCCC 982

RESULT 7
BX463176 1084 bp mRNA linear EST 06-MAY-2004
LOCUS BX463176 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM007YN21 5-PRIME, mRNA sequence.
ACCESSION BX463176
VERSION BX463176 GI:47069010
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Molecular Biology: Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1084)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31033492.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
```

Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4338.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS0DM007YG110P1ec=4338.f>.
Location/Qualifiers
1..1084
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007YN21"
/issue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```
ORIGIN
Query Match 83.8%; Score 839.2; DB 5; Length 1084;
Best Local Similarity 97.3%; Pred. No. 3,2e-177;
Matches 858; Conservative 5; Mismatches 18; Indels 1; Gaps 1;

OY 1 ATGTGAGACTGACCTGATGTTCAACGAGCAAGGCTGTGTGAGAGCTGCTGCTGCGCAG 60
DB 123 ATGTGAGACTGACCTGATGTTCAACGAGCAAGGCTGTGTGAGAGCTGCTGCTGCGCAG 182
OY 61 GACCTGACGCTGTGGGCTGTCACTGTGTGCGCTGTGGGCGCTGTGTGGGCGTGCAGTG 120
DB 183 GACCTGACGCTGTGGGCTGTCACTGTGTGCGCTGTGGGCGCTGTGTGGGCGTGCAGTG 242
OY 121 GGCGTGTGCTCAACAGCGCTCTGTGTGTGCTGCGCACTTACACAGCAAGCCAGATGACC 180
DB 243 GGCGTGTGCTCAACAGCGCTCTGTGTGTGCTGCGCACTTACACAGCAAGCCAGATGACC 302
OY 181 ATGCCGAGCTGATCTTTGTCAACATGGCAGTGGCGCTGTGTCTGACAGGCCCTTGACC 240
DB 303 ATGCCGAGCTGATCTTTGTCAACATGGCAGTGGCGCTGTGTCTGACAGGCCCTTGACC 362
OY 241 CCTGTGACCTGTGTGGCCCCCGGAGCTCCCGATGGGCGCTGTGGAGTGTGGCGCGCA 300
DB 363 CCTGTGACCTGTGTGGCCCCCGGAGCTCCCGATGGGCGCTGTGGAGTGTGGCGCGCA 422
OY 301 GTCCAGTGTGACCTGTGGAGTCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCCACC 360
DB 423 GTCCAGTGTGACCTGTGGAGTCCCTTCAATGTGTCTCACTGTGTGGCCATGTACTCCACC 482
OY 361 GCCCTGTGAGCTGTGACCACTACATGAGCGGCACTGCGCGGCACTACATGGCCAGC 420
DB 483 GCCCTGTGAGCTGTGACCACTACATGAGCGGCACTGCGCGGCACTACATGGCCAGC 542
OY 421 GTGTACAACACGCGGCACTGTGTGCGGCTTGTGTGGGATGTGGCGGCTGTGACCACTTC 480
DB 543 GTGTACAACACGCGGCACTGTGTGCGGCTTGTGTGGGATGTGGCGGCTGTGACCACTTC 602
OY 480 CTCTGTGCTGCTTTTACATCTGTGACGCACTGTGTCCACCGCGGCTAGAGTGGCGCA 539
DB 603 CTCTGTGCTGCTTTTACATCTGTGACGCACTGTGTCCACCGCGGCTAGAGTGGCGCA 662
OY 540 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
DB 663 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
OY 600 ACTGAGCACTCTTACGCGCTGTGTGCTACTTCCGCGCTTCCGAGAGAGAGAGAGAGAGAG 659
DB 723 ACTGAGCACTCTTACGCGCTGTGTGCTACTTCCGCGCTTCCGAGAGAGAGAGAGAGAGAG 782
```

Qy	660	GGACCGGGGACACGGGGCCGGCTGTGAGAGCCCTTGGGACACAGAGCTGCTGTGGGACACCGTGTG	719
Db	783	GGACCGGGGACACGGGGCCGGCTGTGAGAGCCCTTGGGACACAGAGGCTGTGGGACACCGTGTG	842
Qy	720	CACGACGATTTGGGCTCTGTGAGAGCCACATATCTATCTGTGCTGTGGGGACACGCTCATCAT	779
Db	843	CACGATTTTGGGCTCTGTGAGAGCCACATATCTATCTGTGCTGTGGGGACACAGCTCATCAT	902
Qy	780	CTCGGAGGGAAGCCCGGTGACGACGACACTACTTGGGGGCTACTGCACTTTGTGAAGATTT	839
Db	903	CTCGGAGGGAAGGCGCTGTGACGACGACACTACTTGGGGGCTACTGCACTTTGTGAAGATTT	962
Qy	840	CTCCAAACTCTCGGCTCTTCTCCAGACAGCTTTGTGACACCACT	881
Db	963	CTCCAAACTCTGTGGCTTCTCCASMRCTTTGTGAACACTTCT	1004
RESULT 8			
LOCUS	BX338900	1001 bp	mRNA linear EST 07-APR-2004
DEFINITION	BX338900 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1065YE04 5-PRIME, mRNA sequence.		
ACCESSION	BX338900		
VERSION	BX338900.2	GI:46268457	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	L.J.W.B., Gruber,C., Jeasee,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	On May 2, 2003 this sequence version replaced gi:30333744.		
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: beget@genoscope.cns.fr, web : www.genoscope.cns.fr			
1st strand cDNA was primed with a NotI-oli50 (dT) primer. Five primers			
end enriched, double-strand cDNA was digested with Not I and cloned			
into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library			
was normalized. Library was constructed by life technologies, a			
division of invitrogen. This sequence belongs to sequence cluster			
4338.f			
For more information about this cluster, see			
http://www.genoscope.cns.fr/cdna?c=CS0D1065BC02QP1&c=4338.f.			
FEATURES			
SOURCE			
1..1001			
Location/Qualifiers			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0D1065YE04"			
/tissue_type="PLACENTA COT 25-NORMALIZED"			
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oli50 (dT)			
primer. Five primers end enriched, double-strand cDNA was			
digested with Not I and Ecor V sites of the pCMVSPORT 6			
vector. Library was normalized."			
sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match 75.5%; Score 757; DB 5; Length 1001;			
Best Local Similarity 95.2%; Pred. No. 8e-159;			
Matches 789; Conservative 20; Mismatches 15; Indels 5; Gaps 3;			
Qy	1	ATGTGAGACTCAGCTGGTTCACAGGACAGAGGCTGTGTGAGAGAGCTGCTGCCTGCAG	60
Db	129	ATGTGAGACTCAGCTGGTTCACAGGACAGAGGCTGTGTGAGAGAGCTGCTGCCTGCAG	188
Qy	61	GACCTGACGCTGGGGCTGTCACTGTGTGCTGCTGGGCTGTGTGTGGGCGTCCAGTG	120
Db	189	GACCTGACGCTGGGGCTGTCACTGTGTGCTGCTGGGCTGTGTGTGGGCGTCCAGTG	248
Qy	121	GGCCTGTGCTACACGCCCTGCTGTGCTGTGGCAACTTACACAGACAGGCCACGATGAC	180

Db 249 GGCCTGTGCTCAACAGCCCTGCTGTGTGTGGCCACTTACAAGCAGCATATACC 308
Qy 181 ATGCCGACGCTGTATCTTTTGTCACATGACAGTGGCAGGCGCTGTGCTCAGCGCCCTGACC 240
Db 309 ATGCCGACGCTGTATCTTTTGTCACATGACAGTGGCAGGCGCTGTGCTCAGCGCCCTGACC 368
Qy 241 CCTGTGCACTCTCTCGGCCCCCCCAGTCTCCCGGTGTGGGCGCTGTGGAAATTGTGGGGGGGAA 300
Db 369 CCGTGTGACCTGTCTCGGCCCCCCCAGTCTCCCGGTGTGGGCGCTGTGGAAATTGTGGGGGGGAA 428
Qy 301 GTTCCACGTTGGCACTGTGACAGATCCCTCTTAATGTGCTCATCTGTGGGCATGTACTCCACC 360
Db 429 GTTCCACGTTGGCACTGTGACAGATCCCTCTTAATGTGCTCATCTGTGGGCATGTACTCCACC 488
Qy 361 GCCCTGTGTGAGCCCTCGAACCACTATCAGAGCGTGTGCACTGCGCGGACCTTACATGTGGCAGC 420
Db 489 GCCCTGTGTGAGCCCTCGAACCACTATCAGAGCGTGTGCACTGCGCGGACCTTACATGTGGCAGC 548
Qy 421 GTGTACAAACACGCGG-CACGTGTGCGGCTTGTGTGGGGTGTGGCGCGCTGTCTGACCACTT 479
Db 549 GTGTACAAACACGCGGACGTCGTGCGGCTTGTGTGGGGTGTGGCGCGCTGTCTGACCACTT 608
Qy 480 CTCCCTGCGTGTCTTCTTACATCTGACAGCAATGTGTCCACCGCGCGCTTAGAGTGTGCGCAA 539
Db 609 CTCCTCGCTGTCTTCTTACATCTGACAGCAATGTGTCCACCGCGCGCTTAGAGTGTGCGCAA 668
Qy 540 GATGCAGAACGACAGAAGCTGTGCGCACGCGCAGCTGTGTTTATCGGCTACGTGTGTCCAGC 599
Db 669 GATGCAGAACGACAGAAGCTGTGCGCACGCGCAGCTGTGTTTATCGGCTACGTGTGTCCAGC 728
Qy 600 ACTGGCCACCTCTTCTACGCGCTGTGTGTCTTCTCCCGGTCCGACAGGAGACACAGCCCTT 659
Db 729 ACTGGCCACCTCTTCTACGCGCTGTGTGTCTTCTCCCGGTCCGACAGGAGACACAGCCCTT 788
Qy 660 GGACCGGGGACACGGGGCGGCTGTGAGCCCTGTGGGACACAGAGCTGTGTGGACACCGTGT 719
Db 789 GGACCGGGGACACGGGGCGGCTGTGAGCCCTGTGGGACACARGTGTGTGTGGACACCGTGT 848
Qy 720 CACGCAAGTTTGGGCTCTGTGACGCCACATACTGTATCTCTGTGTGGGACACAGGTATCAT 779
Db 849 CACGCAAGTTTGGGCTCTGTGAGCG-CAMACTAATYTKATCTKCTGTGGGAGACACAGG--TCAT 904
Qy 780 CTCGCGAGGGAAGCCCTGTGACGACGACACTACTCTGTGGGCTACTGTGACTTT 828
Db 905 MMTCTVSCGAGGAGCCCGKGGMGACACATCTTGGGTTWKTCTTTT 953

RESULT 9
BX382511
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1023)
L1,W.B., Gruber,C., Jeejee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30439943.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, **Web:** www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster 4338.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSOD1081DA080P1ec=4338.f.
Location/Qualifiers

FEATURES

Source

1..1023
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1081YB16"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime and enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 74.1%; Score 742.8; DB 5; Length 1023;
Best Local Similarity 93.2%; Pred. No. 1.2e-155;
Matches 818; Conservative 11; Mismatches 44; Indels 5; Gaps 5;

```

QY 1 ATGTGAGCTGAGCTGTTCAAGGCGACAGGGCTGTGGAGAGCTGCTGCTGCAG 60
DB 98 ATGTGAGCGGACGCGGTTCAACGCGACAGGCGTGTGAGAGCGCGCTGCTGCAG 157
QY 61 GACCTGCACTGGGGCTGTCACTTTGTGCTGCTGGCGCTGTGTGGCGTGCAGT 120
DB 158 GACCTGCACTGGGGCGGCTGTCACTGCGCGCGCGCGCTGTGTGGCGTGCAGT 217
QY 121 GGCCTGTGCTCAACGCGCTGTGTGTGCTGCTCACTCAACAGCAAGGCGCAGATGAC 180
DB 218 GGCCTGTGCTCAACGCGCTGTGTGTGCTGCTCACTCAACAGCAAGGCGCAGATGAC 277
QY 181 ATGCGGAGCTGTACTTTGTCAATGATGAGTGCAGAGGCTGTGTGCTGAGCGCTGGCC 240
DB 278 ATGCGGAGCTGTACTTTGTCAATGATGAGTGCAGAGGCTGTGTGCTGAGCGCTGGCC 337
QY 241 CCTGTGCACTGCTGCGGCGCGCGCGAGTCCCGGTGGCGCTGTGTGAGTGTGGCGGCA 300
DB 338 CCTGTGCACTGCTGCGGCGCGCGCGAGTCCCGGTGGCGCGCTGTGTGAGTGTGGCGGCA 397
QY 301 GTCCACGTGGCACTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 398 GTCCACGTGGCACTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 457
QY 361 GGCCTGTGAGCTGTCAACCTCAATGAGCTGCACTGCGCGGACCTTACATGCGCAGC 420
DB 458 GGCCTGTGAGCTGTCAACCTCAATGAGCTGCACTGCGCGGACCTTACATGCGCAGC 517
QY 421 GTGTACAACACGCGG-CACGTGTGCGGCTTGTGTGGGTGGCGGCTGTGTGACAGCTT 479
DB 518 GTGTACAACACGCGGCGGCGGCTTGTGTGGGTGGCGGCTGTGTGACAGCTT 577
QY 480 CTCTGCTGCTTCTTATCATCTGCAAGCCATGTGTCAACCGCGGCTTACAGTGGCCAA 539
DB 578 CTCTGCTGCTTCTTATCATCTGCAAGCCATGTGTCAACCGCGGCTTACAGTGGCCAA 637
QY 540 GATGCAAGACGAGAGCTGCGGAGCGGACGCTGTGTTCATGGCTACGTGTGTGCAAGC 599
DB 638 GATGCAAGACGAGAGCTGCGGAGCGGACGCTGTGTTCATGGCTACGTGTGTGCAAGC 697
QY 600 ATTGGCCACCTCTTACGGGCTGTGTCTACTCCCGCTGCGGAGGAGAGCAAGCCCT 659
DB 698 TTTGGCCACCTCTTACGGGCTGTGTCTACTCCCGCTGCGGAGGAGAGCAAGCCCT 757
QY 660 GGACCGGAGACAGGCGCGGCTGAGAGCGCTGCGGACACAGGCTGTGTGTGCGGCTGTG 719
DB 758 GGACCGGAGACAGGCGCGGCTGAGAGCTGTGTGCGGACACAGGCTGTGTGTGCGGCTGTG 817
QY 720 CACGCAATTTGGGCTGTGTGAGCGCACTATCTGCTGTGTGTGAGGAGCAAGGATCAT 779
DB 818 CACGCAATTTGGGCTGTGTGAGCG-CACACTATCTGATCTGCTGTGTGTGAGGAGCAAGGATCAT- 875

```

```

QY 780 CTGCGGAGGAGACCGCTGTGACGACACTACTGCGGCTGTGTGAGAGATTT 839
DB 876 CTGCGGAGGAG-ARCCCGGTGACGACACTACTGCGGCTGTGTGAGAGATTT 933
QY 840 CTCGAACCTCTGCTGTCTGCGAGGCTTGTGTGAC 877
DB 934 YTCMAACTCGBSCTTCTCMACAGCTTGTTAACACTC 971

```

RESULT 10

EX332385 907 bp mRNA linear EST 07-APR-2004
EX332385 BX332385 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
LOCUS CNA clone CSODC011YB12 5-PRIME, mRNA sequence.
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
On May 1, 2003 this sequence version replaced gi:30310147.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4338.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSODC011DA060P1ec=4338.f.
Location/Qualifiers

FEATURES

Source

1..907
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC011YB12"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime and enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 66.4%; Score 665; DB 5; Length 907;
Best Local Similarity 96.6%; Pred. No. 3.1e-138;
Matches 745; Conservative 7; Mismatches 12; Indels 7; Gaps 7;

```

QY 1 ATGTGAGCTGAGCTGTTCAAGGCGACAGGGCTGTGGAGAGCTGCTGCTGCAG 60
DB 113 ATGTGAGCTGAGCTGTTCAACGCGACAGGCGTGTGAGAGCTGCTGCTGCAG 172
QY 61 GACCTGCACTGGGGCTGTCACTTTGTGCTGCTGGCGCTGTGTGGCGTGCAGT 120
DB 173 GACCTGCACTGGGGCTGTCACTTTGTGCTGCTGGCGCTGTGTGGCGTGCAGT 232
QY 121 GGCCTGTGCTCAACGCGCTGTGTGTGCTGCTGCTCACTCAACAGCAAGGCGCAGATGAC 180
DB 223 GGCCTGTGCTCAACGCGCTGTGTGTGCTGCTGCTCACTCAACAGCAAGGCGCAGATGAC 292
QY 181 ATGCGGAGCTGTACTTTGTCAATGATGAGTGCAGAGGCTGTGTGCTGAGCGCTGGCC 240
DB 293 ATGCGGAGCTGTACTTTGTCAATGATGAGTGCAGAGGCTGTGTGCTGAGCGCTGGCC 352

```

QY 241 CTTGTGACC-TGCTGGGCCCCCGAGCTCCGGTGGGCGCTGTGAGTGTGGGCGCGCA 299
 |||||
 Db 353 CCGGTGACCATGTCTGGCCCCCGGAGCTCCGGTGGGCGCTGTGAGTGTGGGCGCGCA 412
 |||||
 QY 300 AGTTCACGTGGCACTGCAATGCCCTTCAAATGTCTCTCACTGGTGGCAATGACTCCAC 359
 |||||
 Db 413 AGTTCACGTGGCACTGCAATGCCCTTCAAATGTCTCTCACTGGTGGCAATGACTCCAC 472
 |||||
 QY 360 CGGCTGTGAGCTGTGACCACTACATGAGCGTGTGACCTGCGGAGCT-ACATGGCCA 418
 |||||
 Db 473 CGGCTGTGAGCTGTGACCACTACATGAGCGTGTGACCTGCGGAGCT-ACATGGCCA 532
 |||||
 QY 419 GCGGTGTACACAGCGGCACTGTGGGCTTCCGTGGGAGGAGCGCTGTGACCACT 478
 |||||
 Db 533 GCGGTGTACACAGCGGCACTGTGGGCTTCCGTGGGAGGAGCGCTGTGACCACT 552
 |||||
 QY 479 TCTCTGTGCTCTTCTACATGTGACCACTGTGTCCACCGCGCGCTGAGTGGCCA 538
 |||||
 Db 593 TCTCTGTGCTCTTCTACATGTGACCACTGTGTCCACCGCGCGCTGAGTGGCCA 652
 |||||
 QY 539 AGATGACAAACGCAAGAGCTGCGGACCGCACTGTGTCTCATGTGGTACGTGTGGCA 598
 |||||
 Db 653 AGATGACAAACGCAAGAGCTGCGGACCGCACTGTGTCTCATGTGGTACGTGTGGCA 711
 |||||
 QY 599 CACTGGGCACTCTACGCGCTGTGTCTCTCCGCGTCCGAGGAGGAGACACGCGCC 658
 |||||
 Db 712 CACTGGGCACTCTACGCGCTGTGTCTCTCCGCGTCCGAGGAGGAGACACGCGCC 771
 |||||
 QY 659 TGGACC-GGGACACGGGCGCGCTGTGAGCCCTGTGACACACAGCGCTGTGTGGCACCGT 717
 |||||
 Db 772 TGGACC-GGGACACGGGCGCGCTGTGAGCCCTGTGACACACAGCGCTGTGTGGCACCGT 830
 |||||
 QY 718 TGCACGCACTTTGGGCTCTGTGACCGCACACTATCTGTCTGTGGGCGCAC 768
 |||||
 Db 831 TGCACGCA-CTTGGGCTGTGTGACCGCC-GACTATCTGTCTGTGTGGGCGMAC 879
 |||||

RESULT 11
 AY403441 670 bp DNA linear GSS 12-DEC-2003

LOCUS
 DEFINITION Pan troglodytes HCM1558 gene, VIRUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.

ACCESSION
 AY403441

VERSION
 AY403441.1 GI:39759424

KEYWORDS
 GSS.

SOURCE
 Pan troglodytes (chimpanzee)

ORGANISM
 Pan troglodytes

REFERENCE
 AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 670)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,
 Adams,M.D. and Cargill,M.

Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

Science 302 (5652), 1960-1963 (2003)

JOURNAL
 PUBMED 14671302

REFERENCE
 2 (bases 1 to 670)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,
 Adams,M.D. and Cargill,M.

Adams,M.D. and Cargill,M.
 Direct Submision

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA

Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

Location/Qualifiers

1. 670

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

JOURNAL

gene <1..>670
 /locus_tag="HCM1558"

Query Match 65.2%; Score 653.8; DB 9; Length 670;
 Best Local Similarity 98.2%; Pred. No. 9,6e-136;
 Matches 658; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 332 TGTCTCATCTGTGTGGGCACTGTAATCCACCGCCCTGTGTGAGCTTCCACACTATATGAGC 391
 |||||
 Db 1 TGTCTCATCTGTGTGGGCACTGTAATCCACCGCCCTGTGTGAGCTTCCACACTATATGAGC 60
 |||||
 QY 392 GTGACATGCGCGGAGCACTCATGAGCGGTGTGAAACACGCGGACGTTGTGGGCTTGG 451
 |||||
 Db 61 GTGACATGCGCGGAGCACTCATGAGCGGTGTGAAACACGCGGACGTTGTGGGCTTGG 120
 |||||
 QY 452 TGTGGGCTGTGCGGCTGTGACCACTTCTCTGTGTCTTCTATCATCTGTGAGCCATG 511
 |||||
 Db 121 TGTGGGCTGTGCGGCTGTGACCACTTCTCTGTGTCTTCTATCATCTGTGAGCCATG 180
 |||||
 QY 512 TGTTCACCGCGGCTGTGAGTGTGCGCAAGATGTGAAACGCAAGAGTGTGCGGACCGC 571
 |||||
 Db 181 TGTTCACCGCGGCTGTGAGTGTGCGCAAGATGTGAAACGCAAGAGTGTGCGGACCGC 240
 |||||
 QY 572 TGTGTTCATCGGCTGTGAGTGTGCGCAAGATGTGAAACGCAAGAGTGTGCGGACCGC 631
 |||||
 Db 241 TGTGTTCATCGGCTGTGAGTGTGCGCAAGATGTGAAACGCAAGAGTGTGCGGACCGC 300
 |||||
 QY 632 CCGCGCTGTGCGGAGGAGACACGCGCCCTGTGACCGGACACGCGGCTGTGAGCCCTTGG 691
 |||||
 Db 301 CCGCGCTGTGCGGAGGAGACACGCGCCCTGTGACCGGACACGCGGCTGTGAGCCCTTGG 360
 |||||
 QY 692 CACACAGCTGTGTGTGTGCGCACTGTGTGACCGGAGTGTGTGGGCTGTGTGAGCCCACTATC 751
 |||||
 Db 361 CACACAGCTGTGTGTGTGCGCACTGTGTGACCGGAGTGTGTGGGCTGTGTGAGCCCACTATC 420
 |||||
 QY 752 TGTCTGTGCGGAGGAGACACGCGTGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
 |||||
 Db 421 TGTCTGTGCGGAGGAGACACGCGTGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 |||||
 QY 812 TGGGCTGTGCACTTGTGTGAGGATTTCTTCCAACTCTGTGCGCTTCTGAGGAGCTTGG 871
 |||||
 Db 481 TGGGCTGTGCACTTGTGTGAGGATTTCTTCCAACTCTGTGCGCTTCTGAGGAGCTTGG 540
 |||||
 QY 872 TGCACCACTTCTTCTTACCGCTGTGACCACTGACCACTTCCCGAGCAAGCTTCCAGGCTGA 931
 |||||
 Db 541 TGCACCACTTCTTCTTACCGCTGTGACCACTGACCACTTCCCGAGCAAGCTTCCAGGCTGA 600
 |||||
 QY 932 TGAAGAAAGCTGCGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 991
 |||||
 Db 601 TGAAGAAAGCTGCGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
 |||||
 QY 992 TGCTGTGCGTA 1001
 |||||
 Db 661 TGCTGTGCGTA 670
 |||||

RESULT 12

B1825260

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LHAM1407 row: a column: 15
 High quality sequence stop: 791.
 Location/Qualifiers

FEATURES

source
 1..851
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5163926"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_119"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 64.4%; Score 647; DB 4; Length 851;
 Best Local Similarity 98.4%; Pred. No. 3.3e-134;

Matches 727; Conservative 0; Mismatches 5; Indels 7; Gaps 7;

```

QY 6 GAGCTGAGCTGTTCAACGCGACAGGCGTGTGAGAGACTGCTGCTGCCAGAGACT 65
DB 114 GAGCTGAGCTG-TCAACGCGACAGGCGTGTGAGAGACTGCTGCTGCCAGAGACT 172
QY 66 GCAGCTGGGCTGTCACTGTTTGTGCTGCTGGGCTGTGGTGGCGCTGCCAGTGGGCT 125
DB 173 GCAGCTGGGCTGTCACTGTTTGTGCTGCTGGGCTGTGGTGGCGCTGCCAGTGGGCT 232
QY 126 GTGCTAACAAGCCCGTGTGCTGCGGCAACCTACAGAAAGCCAGCATGACCATGCC 185
DB 233 GTGCTAACAAGCCCGTGTGCTGCGGCAACCTACAGAAAGCCAGCATGACCATGCC 292
QY 186 GGAAGTACTTTTGTCAACATGAGGAGTGGAGGCGCTGTGCTAGCGCCCTGGCCCTGT 245
DB 293 GGAAGTACTTTTGTCAACATGAGGAGTGGAGGCGCTGTGCTAGCGCCCTGGCCCTGT 352
QY 246 GCACCTGTGCGGCGGCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGCGGAAATCCA 305
DB 353 GCACCTGTGCGGCGGCGGAGCTCCCGGTGGGCGCTGTGAGTGTGGGCGGCGGAAATCCA 412
QY 306 CGTGGCACTGAGATCCCTTCAATGTGTCTCACTGATGAGCATGTACTCCAGCGCCT 365
DB 413 CGTGGCACTGAGATCCCTTCAATGTGTCTCACTGATGAGCATGTACTCCAGCGCCT 472
QY 366 GCTGAGCTCGACCACTACATGAGCGTGAAGCTGCGGAGCACTTACATGAGCGAGCTGTA 425
DB 473 GCTGAGCTCGACCACTACATGAGCGTGAAGCTGCGGAGCACTTACATGAGCGAGCTGTA 532
QY 426 CAACACGGGCAAGTGTGCGGCTTGTGTGGGGTGGGCGGCGCTGTGTG-TACAGGCTTTCCT 484
DB 533 CAACACGGGCAAGTGTGCGGCTTGTGTGGGGTGGGCGGCGCTGTGTG-TACAGGCTTTCCT 592
QY 485 CGCTGCTCTTCAACATGAGCACTGTGTCCACCCGC-GCGCTAGAGTGGGCGCAAGATG 543
DB 593 CGCTGCTCTTCAACATGAGCACTGTGTCCACCCGCAGGCTAGAGTGGGCGCAAGATG 652
QY 544 CAGAACGAGAGCTGCGAGCGCAAGCTGTGTTTCACTGAGCTACGTGTGCGCAGACTG 603
  
```

```

DB 653 CAGAACGAGAGCTGCGAGCGCAAGCTGTGTTCATCGGCTACGTGTGCCAGCACTG 712
QY 604 GCCACCTCTTACGGGCTGTGTCTACTCTCCCGGCTGCGAGGAGGAGCAGCGCCCTGAGC 663
DB 713 GCCACCTCTTACGGGCTGTGTCTACTCTCCCGGCTGCGAGGAGGAGCAGCGCCCTGGA 772
QY 664 CGGACACGCGGCGGCTTGA-GCCCTGCGCAC-AAGGCTGCTGG-TGGCACCCTG-TG 719
DB 773 CGGACACGCGGCGGCTTGAAGGCGCTTGGGCAACAGGCTGCTGATGCGCAGCTGTGG 832
QY 720 CACGCAATTTTGGGCTCTGG 738
DB 833 CACGCAATTTTGGGCTCTAG 851
  
```

RESULT 13
 BOJ12494 996 bp mRNA linear EST 16-JUN-2002
 LOCUS BOJ12494

DEFINITION AGENCOURT 8352557 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6277823
 5', mRNA sequence.

ACCESSION BOJ12494.1 GI:21851393

VERSION BOJ12494

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

REFERENCE NIH-MGC <http://mgi.mgi.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LCM2464 row: m column: 24
 High quality sequence stop: 611.
 Location/Qualifiers

FEATURES

source
 1..996
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6277823"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_113"
 /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 64.4%; Score 645.4; DB 5; Length 996;
 Best Local Similarity 96.1%; Pred. No. 7.6e-134;

Matches 683; Conservative 0; Mismatches 26; Indels 2; Gaps 2;

```

QY 1 ATGTGAGCTGCAAGCTGTTCACAGGACAGGCTGTGAGAGAGCTGCTGCTGCCAG 60
DB 151 ATGTGAGCTGCAAGCTGTTCACAGGACAGGCTGTGAGAGAGCTGCTGCTGCCAG 210
QY 61 GACCTGAGCTGGGCTGTCACTGTTGTGCTGCTGGGCTGTGGTGGGCGGCTGCCAGT 120
DB 211 GACCTGAGCTGGGCTGTCACTGTTGTGCTGCTGGGCTGTGGTGGGCGGCTGCCAGT 270
QY 121 GGCCTGTGCTAACAGCCCTGCTGTGTGCGCAACTTACAGCAAGGCAAGCATGACC 180
  
```



```

VERSION      AY403442.1  GI:39759425
KEYWORDS
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              (bases 1 to 996)
              Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
              Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferrieri,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
              Adams,M.D. and Cargill,M.
              Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
              Science 302 (5652), 1960-1963 (2003)
JOURNAL      14671302
FUEMED       2 (bases 1 to 996)
REFERENCE    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
AUTHORS      Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferrieri,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
              Adams,M.D. and Cargill,M.
              Direct Submision
              Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
              This sequence was made by sequencing genomic exons and ordering
              them based on alignment:
FEATURES
  source      location/Qualifiers
              1..996
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"
              <1..>996
              /locus_tag="HCM1558"
ORIGIN
Query Match 60.9%; Score 610; DB 9; Length 996;
Beet Local Similarity 76.9%; Pred. No. 6,3e-126;
Matches 771; Conservative 0; Mismatches 225; Indels 6; Gaps 2;
QY 1 ATGTGAGAGCTGACAGGCTTTCACAGGACAGGAGAGAGAGTGCCTGCTGCAG 60
DB 1 ATGTGAGAGCTGACAGGCTTTCACAGGACAGGAGAGAGAGTGCCTGCTGCAG 60
QY 61 GACCTGACAGTGGGGCTGTGACATGTTGTGCTGCTGAGGCTGGTGGTGGGCGTGGCAATG 120
DB 58 AACCTGGCGCTGGGGCTGTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 117
QY 121 GGCCTGTGCTAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 118 AGCTTAGAGCTTATATGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 177
QY 181 ATGCCGAGCTGTACTTGTTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 178 ATGCCGAGCTGTACTTGTTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 237
QY 241 CCTGTGACAGTGTGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 300
DB 238 CTTGCTTACCTGTGTGGCCCTGCGCCCACTCAAGTGGCCCTGTGAGGCTCAGCAGTGA 297
QY 301 GTTCAAGTGGCACTGAGATCCCTTCAATGATGCTCACTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 298 GCCCATGTGACACTGCTATCTCTTCAACGTGGCTTCCCTGATGACATGATCTCCACT 357
QY 361 GGCCTGTGAGGCTGACACCTAATGAGGCTGAGCTGCGCGAGCACTAATGAGGAGC 420
DB 358 GCACTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT 417
QY 421 GTGTACAACAGCGGCACTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT 480
DB 418 GTGTACAACAGCGGCACTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT 477
QY 481 TCTTGGCTGCTTCTTCAATGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT 540
DB 478 TCTTCCGCTCTTCTTCAATGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT 537

```

```

QY 541 ATGTGAGAGCGAAGAGTGCAGCGGACAGCTGTGTTGATCGGCTAAGTGGAGCAGCA 600
DB 538 ATGTGAGAGCGAAGAGTGCAGCGGACAGCTGTGTTGATCGGCTAAGTGGAGCAGCA 597
QY 601 CTGGCCACCTCTTACCGGCTGAGCTACTCTCCCGCTCCGAGGAGAGACAGCCCTG 660
DB 598 CTGGCTGTGTGTATGCTCTTGGACATCTGTGAATCGGAGAGAGACACACCCCTG 657
QY 661 GATCGGAGACAGGCGCGGCTGAGAGCTTGTGAGCAGACAGGCTGCTGTGAGCAGGCTGTC 720
DB 658 GATCGGAGACAGGCGCGGCTGAGAGCTTGTGAGCAGACAGGCTGCTGTGAGCAGGCTGTC 717
QY 721 AGCAGATTGGGCTCTGAGACGGCACAATGATCTGCTGGGAGACAGGCTATCATC 780
DB 718 ACTCATTTGGCTCTGAGACGGCACAATGATCTGCTGGGAGACAGGCTATCATC 774
QY 781 TCGCAGAGGAGACCGGCTGAGACGACACTACCTGGGCTTACGACTTGTGAGAGATTC 840
DB 775 TCACGGGGAGAGACCGGCTGAGAGGCTATGAGGCTTACGAGGCTTGTGAGAGATTC 834
QY 841 TCCAAATCTCGGCTTCTTCCAGAGCTTGTGACACCACTTCTACCGCTACATGAG 900
DB 835 GCGAAGTCTGGCTTCTTCAAGCAGTCTGTGACACCACTTCTACCGCTTACATGAG 894
QY 901 CAGAGCTTCCCGAGAGCTCCAAAGGCTGATGAAAAAGCTGCGCTGGGAGACCGGAC 960
DB 895 AAAGCTTCCCGGAGAGCTCCCGGCTGATGAAAAAGCTGCGCTGGGAGACCGGAC 954
QY 961 TGTCTCCCGGAGCAGATGGGGGTGAGCAGAGTGTGCGTGA 1002
DB 955 TGTCTCCCGGAGCAGCTTGGGAGATGAGCAGATGATGAGCAG 996

```

Search completed: December 15, 2004, 02:34:59
Job time : 3160.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 18:35:19 / Search time 53.3965 Seconds
(without alignments)
2237.166 Million cell updates/sec

Title: US-09-995-225b-2

Perfect score: 1733
Sequence: 1 MMSCSWFNGTGLVBLPACQ.....LPCDRHCSPEHMKVQVLA 333

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_23Sep04:.*
1: geneseq1980s:.*
2: geneseq1990s:.*
3: geneseq2000s:.*
4: geneseq2001s:.*
5: geneseq2002s:.*
6: geneseq2003as:.*
7: geneseq2003bs:.*
8: geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1733	100.0	333	4	AAU25607	Aau25607 Human G P
2	1733	100.0	333	5	ABJ04068	Abj04068 Human G P
3	1733	100.0	333	5	AAU97734	Aau97734 Human TGR
4	1733	100.0	333	6	ABR56303	Abt56303 Human G-P
5	1733	100.0	333	6	ABU09900	Abu09900 Human G-P
6	1733	100.0	333	6	AdC86563	AdC86563 Human GPC
7	1733	100.0	333	8	AD029606	Ad029606 Human GPC
8	1733	100.0	370	4	AAW25659	Aaw25659 Human pro
9	1733	100.0	695	6	ABR56305	Abt56305 pc901HISG
10	1733	100.0	714	6	ABR56304	Abt56304 pc901HISG
11	1733	100.0	735	6	ABR56306	Abt56306 pc901HISG
12	1726	99.6	333	3	AAV57285	Aav57285 Human GPC
13	1726	99.6	333	4	AAU25584	Aau25584 Human G P
14	1726	99.6	333	5	ABU54867	Abu54867 Human G-P
15	1717	99.1	333	3	AAI78856	Aai78856 Human che
16	1717	99.1	333	6	ABP81753	Abp81753 Human G P
17	1276	73.6	333	8	ABD029607	Abd029607 Mouse GPC
18	795	45.9	153	5	ABBO4632	Abbo4632 Human ops
19	711	41.0	142	5	ABP10980	Abp10980 Human ORF
20	271	15.6	375	2	AAW03739	Aaw03739 rchd523 g
21	271	15.6	375	2	AAW89302	Aaw89302 Human rch
22	271	15.6	375	2	AAW95158	Aaw95158 Human act
23	271	15.6	375	3	AAAB19625	Aab19625 Human rch
24	271	15.6	375	3	AAV78506	Aav78506 Protein s
25	271	15.6	375	3	AAV68445	Aav68445 Rchd523 a

ALIGNMENTS

RESULT 1	AAU25607	standard; protein; 333 AA.
ID	AAU25607	
XX	AAU25607;	
AC	18-DEC-2001	(first entry)
DT	18-DEC-2001	(first entry)
XX	Human G Protein-Coupled Receptor (GPCR) polypeptide #54.	
XX	Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia; attention deficit disorder; anxiety; depression; bipolar disorder;	
KW	neurological disorder; Huntington's disease; dementia; obesity; anorexia;	
KW	metabolic disorder; Parkinson's disease; Tourette's syndrome; chromosom;	
KW	type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;	
KW	cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;	
KW	viral infection; immunostimulant; neuroleptic; nootropic; tranquilliser;	
KW	antidepressant; anorectic; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200162797-A2.	
XX		
PD	30-AUG-2001.	
PF	23-FEB-2001; 2001WO-US005676.	
XX		
XX	23-FEB-2000; 2000US-0184247P.	
PR	23-FEB-2000; 2000US-0184303P.	
PR	23-FEB-2000; 2000US-0184304P.	
PR	23-FEB-2000; 2000US-0184305P.	
PR	23-FEB-2000; 2000US-0184397P.	
PR	02-MAR-2000; 2000US-0186457P.	
PR	03-MAR-2000; 2000US-0186810P.	
PR	09-MAR-2000; 2000US-0188064P.	
PR	13-MAR-2000; 2000US-0188809P.	
PR	03-APR-2000; 2000US-0194344P.	
PR	23-JUN-2000; 2000US-0213861P.	
PR	11-JUL-2000; 2000US-0217369P.	
PR	11-JUL-2000; 2000US-0217370P.	
PR	14-JUL-2000; 2000US-0218337P.	
PR	20-JUL-2000; 2000US-0218492P.	
XX		
PA	(PMAA) PHARMACIA & UPJOHN CO.	
XX		
PI	Vogel G, Wood LS, Parodi LA, Lind P,	
XX	WPL; 2001-570628/64.	
DR		

26	271	15.6	375	7	AD110059	Ad110059 Novel hum
27	271	15.6	375	8	ADFP1255	Adfp1255 Human end
28	269	15.5	375	8	AAV90658	Aav90658 Human mut
29	269	15.5	375	3	AAV90652	Aav90652 Human mut
30	269	15.5	375	7	ADCC2713	Adc22713 Human G P
31	269	15.5	375	7	ADCC2721	Adc22721 Human G P
32	269	15.5	375	7	ADH14194	Adh14194 Mutated h
33	269	15.5	375	7	ADH14186	Adh14186 Mutated h
34	267	15.4	375	2	AAW22338	Aaw22338 PEG-1 pro
35	267	15.4	375	3	AAV59177	Aav59177 Human CEP
36	267	15.4	375	3	AAV90628	Aav90628 Human G P
37	267	15.4	375	3	AAV90624	Aav90624 Human G P
38	267	15.4	375	6	ABG73996	Abg73996 Human CMK
39	267	15.4	375	6	ABU08546	Abu08546 Human hep
40	267	15.4	375	6	ABP81804	Abp81804 Human G P
41	267	15.4	375	7	ADCC2589	Adc22589 Human G P
42	267	15.4	375	7	ADCC2575	Adc22575 Human G P
43	267	15.4	375	7	ADH14048	Adh14048 Human GPR
44	267	15.4	375	7	ADH14062	Adh14062 Human CEP
45	267	15.4	375	8	AD029390	Ad029390 Human GPC

DR N-PSDB; AAS42859.

XX New isolated nucleic acid encoding a new G-protein coupled receptor
PT polypeptide for detecting receptor modulators that can treat mental
PT disorders, such as schizophrenia, anxiety, depression, or obesity.

XX Claim 35; Page 91; 279pp; English.

XX Sequences AAU25554-AAU25616 represent human G-protein coupled receptor
CC (GPCR) polypeptides of the invention. The proteins and their associated
CC DNA sequences can be used to identify compounds which bind to GPCR
CC polypeptides and in screening for compounds that modulate GPCR activity.
CC By screening a human subject for the presence of mutations in GPCR DNA, a
CC GPCR-related disorder or a genetic predisposition can be diagnosed. The
CC sequences can also be used for treatment and prevention of mental
CC disorders such as schizophrenia, attention deficit disorder, anxiety,
CC depression, dementia and bipolar disorder, neurological disorders such as
CC Huntington's disease, Parkinson's disease and Tourette's syndrome,
CC metabolic disorders such as obesity, anorexia and type 2 diabetes,
CC cardiovascular disorders such as thrombosis, myocardial infarction,
CC cardiomyopathy and atherosclerosis, viral infections caused by HIV and
CC cancers

XX Sequence 333 AA;

Query Match 100.0%; Score 1733; DB 4; Length 333;

Best Local Similarity 100.0%; Pred. No. 4e-177; Mismatches 0; Gaps 0;

Matches 333; Conservative 0; Indels 0; Gaps 0;

QY 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLGVGVPGVGCYNALLVLANHSHKASMT 60
DB 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLGVGVPGVGCYNALLVLANHSHKASMT 60

QY 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120

QY 121 ALISLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTSFSSLLFYICSHVSTRALBEC 180
DB 121 ALISLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTSFSSLLFYICSHVSTRALBEC 180

QY 181 MONEAADATLVFIIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHRLLVATVC 240
DB 181 MONEAADATLVFIIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHRLLVATVC 240

QY 241 TOFGIWTPHYLLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSVTPPLRYNM 300
DB 241 TOFGIWTPHYLLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSVTPPLRYNM 300

QY 301 QSFPSKLORLMKKLPCCDRHSCPDHMGVQOYLA 333
DB 301 QSFPSKLORLMKKLPCCDRHSCPDHMGVQOYLA 333

RESULT 2
ABJ04068
ID ABJ04068 standard; protein; 333 AA.

XX ABJ04068;
XX
XX 11-OCT-2002 (first entry)

XX Human G protein coupled receptor hrUP28.

XX Human; G-protein coupled receptor; GPCR; hrUP28; hrUP29; hrUP30; hrUP31;
XX hrUP32; hrUP33; hrUP34; hrUP35; hrUP36; hrUP37.

XX Homo sapiens.

XX WO200242461-A2.

XX 30-MAY-2002.

XX

PF 26-NOV-2001; 2001WO-US044386.

XX 27-NOV-2000; 2000US-0253404P.

PR 12-DEC-2000; 2000US-0255366P.

PR 20-FEB-2001; 2001US-0270266P.

PR 20-FEB-2001; 2001US-0270266P.

PR 06-APR-2001; 2001US-0282032P.

PR 06-APR-2001; 2001US-0282358P.

PR 06-APR-2001; 2001US-0282358P.

PR 14-MAY-2001; 2001US-0290917P.

PR 31-JUL-2001; 2001US-0309208P.

XX (AREN-) ARENA PHARM INC.

XX Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;

DR WPI: 2002-566565/60.

DR N-PSDB; ABT04866.

XX Novel endogenous and non-endogenous versions of G protein-coupled

XX receptor useful for identification of candidate compounds as receptor agonists or inverse

XX agonists or antagonists for use as therapeutic agents. The present sequence

XX is a GPCR protein of the invention

XX Sequence 333 AA;

Query Match 100.0%; Score 1733; DB 5; Length 333;

Best Local Similarity 100.0%; Pred. No. 4e-177; Mismatches 0; Gaps 0;

Matches 333; Conservative 0; Indels 0; Gaps 0;

QY 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLGVGVPGVGCYNALLVLANHSHKASMT 60
DB 1 MMSGWFNGTGLVEELPACODLQGLSLSLGLGVGVPGVGCYNALLVLANHSHKASMT 60

QY 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120

QY 121 ALISLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTSFSSLLFYICSHVSTRALBEC 180
DB 121 ALISLDHYIERALPRTYMASVYNTNRHVCGFVWGALLTSFSSLLFYICSHVSTRALBEC 180

QY 181 MONEAADATLVFIIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHRLLVATVC 240
DB 181 MONEAADATLVFIIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHRLLVATVC 240

QY 241 TOFGIWTPHYLLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSVTPPLRYNM 300
DB 241 TOFGIWTPHYLLILGHTVLIISRGKPVDAHYLGILHFPVDFSKLIAFSSSVTPPLRYNM 300

QY 301 QSFPSKLORLMKKLPCCDRHSCPDHMGVQOYLA 333
DB 301 QSFPSKLORLMKKLPCCDRHSCPDHMGVQOYLA 333

RESULT 3
AAU97734
ID AAU97734 standard; protein; 333 AA.

XX AAU97734;

XX 27-AUG-2002 (first entry)

XX Human TGR339 polypeptide.

XX

XX

KW Human; TGR339; receptor; G-protein coupled receptor; GPCR; TGR;
 KW TGR-associated disorder; signal transduction; renal failure; nephritis;
 KW hypochyroidism; hypogonadism; retinitis pigmentosa; growth disorder;
 KW diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;
 KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;
 KW circadian rhythm.
 XX
 OS Homo sapiens.
 XX
 PN MO200242458-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 21-NOV-2001; 2001MO-US043404.
 XX
 PR 22-NOV-2000; 2000US-0252841P.
 PR 22-DEC-2000; 2000US-0257636P.
 PR 12-JAN-2001; 2001US-0261377P.
 PR 28-MAR-2001; 2001US-0279554P.
 PR 29-MAR-2001; 2001US-0280696P.
 XX
 PA (TULSA-) TULARIK INC.
 XX
 PI Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupta JS;
 XX
 DR WPI; 2002-463633/49.
 DR N-PSDB; ABR6286.
 XX
 PT New isolated G-protein couple receptor polypeptide, termed TGR, for
 PT diagnosis and treatment of diseases such as renal failure, nephritis,
 PT hypochyroidism, diabetes insipidus, and disturbances of thirst and sleep.
 XX
 PS Claim 33; Page 74; 98pp; English.
 XX
 XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),
 CC termed TGR, and its associated nucleic acid. The sequences of the
 CC invention are useful for identifying a compound that modulates signal
 CC transduction and for identifying a mammal having a TGR-associated
 CC disorder. The proteins and nucleic acids are useful in diagnosis and
 CC treatment of diseases or conditions such as renal failure, nephritis,
 CC hypochyroidism, hypogonadism, retinitis pigmentosa, growth disorders,
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,
 CC sleep, temperature regulation, appetite, blood pressure or any other
 CC syndrome or disease associated with the hypothalamus. The sequences can
 CC be used in regulation of circadian rhythms, for use as genetic markers
 CC for the identification of particular diseases associated with diseases resulting
 CC from GPCR inactivation in particular cell types and for identification of
 CC modulators of GPCR signal transduction. This sequence represents the
 CC human TGR339 polypeptide
 CC
 XX
 SQ Sequence 333 AA;
 Query Match 100.0%; Score 1733; DB 5; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MMSCSWNGTGLVEEIPACODLQGLSLSLGLVGVPGVGLCYNALLVLANIHSKASMT 60
 DB 1 MMSCSWNGTGLVEEIPACODLQGLSLSLGLVGVPGVGLCYNALLVLANIHSKASMT 60
 QY 61 MPVVYFVNMAVAGLVISALAPVHLGPSSRMALMSVGEVHALQIPFNVSSLVAMYST 120
 DB 61 MPVVYFVNMAVAGLVISALAPVHLGPSSRMALMSVGEVHALQIPFNVSSLVAMYST 120
 QY 121 ALISLDHYIERALPRTYMASVYNTTRHVCQFVWGALLTSFSSILFYICSHVSTRALBCK 180
 DB 121 ALISLDHYIERALPRTYMASVYNTTRHVCQFVWGALLTSFSSILFYICSHVSTRALBCK 180
 QY 181 MONAADAATLVETIGVVPALATLYALVLLSRVREDTPLDRDTGRLEPSAHLVAVTC 240
 DB 181 MONAADAATLVETIGVVPALATLYALVLLSRVREDTPLDRDTGRLEPSAHLVAVTC 240
 QY 241 TQGLMTPHYLILIGHTVILSRGKPVDAHYLGHLHPKDFSKLAPSSSPFTPLLYRYNM 300

DB 241 TQGLMTPHYLILIGHTVILSRGKPVDAHYLGHLHPKDFSKLAPSSSPFTPLLYRYNM 300
 QY 301 QSEFSPKQRLMKKLPGBDRHCSBDMGVQOVLTA 333
 DB 301 QSEFSPKQRLMKKLPGBDRHCSBDMGVQOVLTA 333
 RESULT 4
 ABR56303
 ID ABR56303 standard; protein; 333 AA.
 XX
 AC ABR56303;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 XX Human G-protein coupled receptor 901.
 DE Human G-protein coupled receptor 901.
 XX
 KW Human; anorectic; antidiabetic; antilipemic; hypothalamus;
 KW G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
 KW cibophobia; anorexia nervosa.
 XX
 OS Homo sapiens.
 XX
 PN WO2003030936-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 02-OCT-2002; 2002MO-JP010250.
 PF 02-OCT-2001; 2001JP-00306872.
 PR 02-OCT-2001; 2001JP-00306872.
 XX
 PA (SUMU) SUMITOMO PHARM CO LTD.
 XX
 PI Suguru E, Tsuchida A, Yamanaka M, Taiji M;
 XX
 DR WPI; 2003-354886/33.
 DR N-PSDB; ACCT0841.
 XX
 PT Inhibitors of expression or activity of G-protein coupled receptor 901
 PT for treatment of lifestyle-related diseases and cibophobia.
 XX
 XX Claim 1; Page 64-65; 91pp; Japanese.
 XX
 CC The present invention relates to novel remedies for the treatment of
 CC diseases containing as an active component an inhibitor of the expression
 CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
 CC for treatment of cibophobia containing as an active component a
 CC potentiator of the expression or activity of G-protein coupled receptor
 CC 901. The diseases which can be treated include obesity, diabetes and
 CC hyperlipaemia, and cibophobia (anorexia nervosa). The present human
 CC sequence is G-protein coupled receptor 901
 CC
 XX
 SQ Sequence 333 AA;
 Query Match 100.0%; Score 1733; DB 6; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MMSCSWNGTGLVEEIPACODLQGLSLSLGLVGVPGVGLCYNALLVLANIHSKASMT 60
 DB 1 MMSCSWNGTGLVEEIPACODLQGLSLSLGLVGVPGVGLCYNALLVLANIHSKASMT 60
 QY 61 MPVVYFVNMAVAGLVISALAPVHLGPSSRMALMSVGEVHALQIPFNVSSLVAMYST 120
 DB 61 MPVVYFVNMAVAGLVISALAPVHLGPSSRMALMSVGEVHALQIPFNVSSLVAMYST 120
 QY 121 ALISLDHYIERALPRTYMASVYNTTRHVCQFVWGALLTSFSSILFYICSHVSTRALBCK 180
 DB 121 ALISLDHYIERALPRTYMASVYNTTRHVCQFVWGALLTSFSSILFYICSHVSTRALBCK 180
 QY 181 MONAADAATLVETIGVVPALATLYALVLLSRVREDTPLDRDTGRLEPSAHLVAVTC 240
 DB 181 MONAADAATLVETIGVVPALATLYALVLLSRVREDTPLDRDTGRLEPSAHLVAVTC 240

Db 181 MONAADAATLVFIGVVPALATYALVLSRVREDDTPLDRDTGRLSPSAHRLVAATVC 240
 QY 241 TQFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFKVDFSKSLAFSSFPVPLRYNN 300
 Db 241 TQFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFKVDFSKSLAFSSFPVPLRYNN 300
 QY 301 QSPFSKIQRLMKKLPDGDHRCSPDHMGVOQVLA 333
 Db 301 QSPFSKIQRLMKKLPDGDHRCSPDHMGVOQVLA 333

RESULT 5

ABU09900 standard; protein; 333 AA.

AC ABU09900;

DT 11-AUG-2003 (first entry)

DE Human G-protein coupled receptor 23992.

KW Human; receptor; G-protein coupled receptor; gene therapy; pneumonia;
 KW immune disorder; Crohn's disease; Grave's disease; respiratory disorder;
 KW diphtheria; haematopoietic disorder; leukaemia; systemic sclerosis;
 KW prostatic disorder; benign prostate hypertrophy; tumour; breast disease;
 KW acute mastitis; Paget's disease; muscular disorder; rhabdomyosarcoma;
 KW neurological disorder; cerebral oedema; Parkinson's disease; atrophy;
 KW blood vessel disorder; atherosclerosis; testicular disease; syphilis;
 KW kidney disorder; cystic renal dysplasia; hyperthyroidism; cretinism; AIDS;
 KW cardiovascular disease; heart failure; pericarditis; pancreatic disease;
 KW pancreatitis; diabetes mellitus; thymus disease; thymic hypoplasia;
 KW Hodgkin disease; spleen disease; splenomegaly; Gaucher disease;
 KW liver disease; hepatic failure; alcoholic liver disease; T-cell disorder;
 KW systemic lupus erythematosus; B-cell disorder; Burkitt lymphoma;
 KW multiple myeloma; platelet disorder; thrombocytopenia;
 KW haemolytic-uraemic syndrome.

OS Homo sapiens.

PN US2003017539-A1.

PD 23-JAN-2003.

EF 07-JUN-2002; 2002US-00165844.

PR 02-JUN-1998; 98US-00088857.
 PR 02-SEP-1998; 98US-00145745.
 PR 21-JUN-1999; 99US-00234923.
 PR 02-JUN-1999; 99US-00324465.
 PR 28-JUN-1999; 99US-00340880.
 PR 26-AUG-1999; 99US-00383745.
 PR 16-DEC-1999; 99US-00464685.
 PR 18-DEC-2000; 2000US-00741783.

PA (MILL-) MILLENNIUM PHARM INC.

PI Glucksmann MA, Hodge MR, Hunter JJ, Rudolph-Owen LA;
 PI Siles-Santiago I, Welch NS;

DR MPI; 2003-401672/38.

DR N-PSDB; ACA61176.

PT New nucleic acid molecule encoding a G-protein coupled receptor
 PT polypeptide, e.g. 2871, 14926, 17723 or 23992, useful for diagnosing
 PT and/or treating cancer or immune, respiratory, hematologic or
 PT cardiovascular disorders.

PS Claim 9; Fig 21; 149pp; English.

CC The invention relates to an isolated G-protein coupled receptor nucleic
 CC acid molecule encoding the G-protein coupled receptors 2871, 14926, 17723
 CC and 23992. The nucleic acid molecule is useful in monitoring, diagnosing

CC and treating immune disorders (e.g. Crohn's disease and Grave's disease);
 CC respiratory disorders (e.g. pneumonia or diphtheria); haematopoietic
 CC disorders (e.g. leukaemia and systemic sclerosis); prostatic disorders
 CC (e.g. benign prostate hypertrophy and tumours); breast diseases (e.g.
 CC acute mastitis and Paget's disease); muscular disorders (e.g.
 CC rhabdomyosarcoma); neurological disorders (e.g. cerebral oedema and
 CC Parkinson's disease); disorders involving blood vessels (e.g.
 CC atherosclerosis); diseases involving testis and epididymis (e.g.
 CC kidney disorders (e.g. hyperthyroidism and cretinism);
 CC cardiovascular diseases (e.g. heart failure and pericarditis); pancreatic
 CC diseases (e.g. pancreatitis and diabetes mellitus); diseases involving
 CC the thymus (e.g. thymic hypoplasia and Hodgkin disease); diseases
 CC involving the spleen (e.g. splenomegaly and Gaucher disease); diseases
 CC involving the liver (e.g. hepatic failure and alcoholic liver disease);
 CC -cell disorders (e.g. systemic lupus erythematosus and AIDS); T
 CC disorder (e.g. Burkitt lymphoma and multiple myeloma); platelet disorders
 CC (e.g. thrombocytopenia and haemolytic-uraemic syndrome). The nucleic
 CC acids may also be used in chromosome mapping, tissue typing,
 CC pharmacogenomics and forensic biology, and as surrogate markers. The
 CC present sequence represents the amino acid sequence of the human G-
 CC protein coupled receptor 23992

SQ Sequence 333 AA;

Query Match 100.0%; Score 1733; DB 6; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWSGWFNGTGVLEELPACODLQGLSLSLGLVGVGPGLCNMLVLANHSKASMT 60
 Db 1 MWSGWFNGTGVLEELPACODLQGLSLSLGLVGVGPGLCNMLVLANHSKASMT 60
 QY 61 MPDYFVMAVAGVLSALAPVHLGPPSSRAKMSVGEVHALQIPNVSSLVAMYST 120
 Db 61 MPDYFVMAVAGVLSALAPVHLGPPSSRAKMSVGEVHALQIPNVSSLVAMYST 120
 QY 121 ALSLDHYIERALPRTYMASVNTTRHVGCFVWGALITFSLSLPTICSHVSTRALCKK 180
 Db 121 ALSLDHYIERALPRTYMASVNTTRHVGCFVWGALITFSLSLPTICSHVSTRALCKK 180
 QY 181 MONAADAATLVFIGVVPALATYALVLSRVREDDTPLDRDTGRLSPSAHRLVAATVC 240
 Db 181 MONAADAATLVFIGVVPALATYALVLSRVREDDTPLDRDTGRLSPSAHRLVAATVC 240
 QY 241 TQFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFKVDFSKSLAFSSFPVPLRYNN 300
 Db 241 TQFGMTPHYLILGHTVLIISRGKPVDAHYLGILHFKVDFSKSLAFSSFPVPLRYNN 300
 QY 301 QSPFSKIQRLMKKLPDGDHRCSPDHMGVOQVLA 333
 Db 301 QSPFSKIQRLMKKLPDGDHRCSPDHMGVOQVLA 333

RESULT 6

ADCB6563
 ID ADCB6563 standard; protein; 333 AA.

AC ADCB6563;

DT 01-JAN-2004 (first entry)

DE Human GPCR protein SEQ ID NO:1016.

DE human; GPCR; guanosine triphosphate-binding protein, coupled receptor;
 KW gene therapy.

OS Homo sapiens.

PN EP1270724-A2.

PD 02-JAN-2003.

PF 18-JUN-2002; 2002EP-00013517.
 XX
 XX 18-JUN-2001; 2001JP-00246789.
 XX
 PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATORIO.
 XX
 PI Suwa M, Asai K, Akiyama Y, Aburatani H;
 XX
 XX WPI, 2003-315783/31.
 DR N-PSDB; ADC66562.
 XX
 XX New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX
 PS Claim 2; SEQ ID NO 1016; 28pp; English.
 XX
 CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC65549-ADC67617 represent GPCR's of the invention.
 XX
 SQ Sequence 333 AA;
 Query Match 100.0%; Score 1733; DB 7; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MWSGSMWNGTGLVEELPACQDLQGLSLSLGLVGVPGVGLCYNALLVLANHKSASMT 60
 DB 1 MWSGSMWNGTGLVEELPACQDLQGLSLSLGLVGVPGVGLCYNALLVLANHKSASMT 60
 QY 61 MPDYVFNNMVAAGLVLSALAPVHLGPPSSRMALMSVGEVHVLAOLPNNVSLVAMYST 120
 DB 61 MPDYVFNNMVAAGLVLSALAPVHLGPPSSRMALMSVGEVHVLAOLPNNVSLVAMYST 120
 QY 121 ALLSLDHYIRALPRTYMASVYTRHVGCFVWGALLTSFSSLLFYICSHVSTRALBECAR 180
 DB 121 ALLSLDHYIRALPRTYMASVYTRHVGCFVWGALLTSFSSLLFYICSHVSTRALBECAR 180
 QY 181 MOWAEADATLVPIGTVVPAATLYALVLSRVARETPIIDRTGRLPEPSAHLIVATVC 240
 DB 181 MOWAEADATLVPIGTVVPAATLYALVLSRVARETPIIDRTGRLPEPSAHLIVATVC 240
 QY 241 TORGLMTPHYLLILGHVTIISRGKPVNAHYLGILHFKDPSKILAFSSSVTPILLYRYNN 300
 DB 241 TORGLMTPHYLLILGHVTIISRGKPVNAHYLGILHFKDPSKILAFSSSVTPILLYRYNN 300
 QY 301 QSEFPSKIQRLMKKLPCGDRHCSPDHMGVQVLA 333
 DB 301 QSEFPSKIQRLMKKLPCGDRHCSPDHMGVQVLA 333
 RESULT 7
 ADO29606
 ID ADO29606 standard; protein; 333 AA.
 XX
 AC ADO29606;
 XX
 XX 29-JUL-2004 (first entry)
 DT
 XX Human GPCR PGR8, SEQ ID NO:708.
 XX
 XX G proteain-coupled receptor; GPCR; drug screening; diagnosis;
 KM transgenic mouse; neurological disorder; adrenal gland disorder;
 KM colon disorder; intestinal disorder; cardiovascular disorder;
 KM muscular disorder; blood disorder; immune disorder; bone disorder;
 KM joint disorder; metabolic disorder; nutritive disorder; cancer;
 KM kidney disorder; liver disorder; lung disorder; breast disorder;

KM ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KM skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KM thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
 KM cytostatic; antiinflammatory; vasotropic; antiangiogenic; antiallergic;
 KM CNS; central nervous system; respiratory; antidiabetic; antidiabetic;
 KM virucide; hepatotropic; antibacterial; antianemic; antiseborrheic;
 KM dermatological; antiulcer; antihypertensive; anorectic;
 KM immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KM receptor.
 XX
 XX Homo sapiens.
 OS
 XX
 XX W02004040000-A2.
 PN
 XX 13-MAY-2004.
 PD
 XX
 XX 09-SEP-2003; 2003MO-US028226.
 PE
 XX
 XX 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 XX (PRIM-) PRIMAL INC.
 PA
 PI Galtanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Medisen L, McIlwain KL, Pavlova MN, Vasiliadis D, Zeng H;
 XX
 XX WPI, 2004-390329/36.
 DR N-PSDB; ADO30022.
 XX
 XX Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 PS Claim 151; SEQ ID NO 708; 542pp; English.
 XX
 CC The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders (e.g., autoimmune disorders or
 CC anemias or leukemias); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_poc_sequences.
 XX
 SQ Sequence 333 AA;
 Query Match 100.0%; Score 1733; DB 8; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4e-177;
 Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 17-APR-2003.
PD 02-OCT-2002; 2002WO-JP010250.
PF 02-OCT-2001; 2001JP-00306872.
PR 02-OCT-2001; 2001JP-00306872.
XX (SUMU) SUMITOMO PHARM CO LTD.
PA Suguru E, Tsuchida A, Yamana M, Taiji M;
PI WPI; 2003-354886/33.
DR N-PSDB; ACC70860.
XX
PT Inhibitors of expression or activity of G-protein coupled receptor 901
for treatment of lifestyle-related diseases and cibophobia.
XX
PS Disclosure; Page 79-81; 91pp; Japanese.
XX The present invention relates to novel remedies for the treatment of
diseases containing as an active component an inhibitor of the expression
or activity of hypothalamus-expressed G-protein coupled receptor 901 and
CC for treatment of cibophobia containing as an active component a
CC potentiator of the expression or activity of G-protein coupled receptor
901. The diseases which can be treated include obesity, diabetes and
CC hyperlipemia, and cibophobia (anorexia nervosa). The present sequence
was used to illustrate the invention
XX
SQ Sequence 695 AA;

Query Match 100.0%; Score 1733; DB 6; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.1e-176; Indels 0; Gaps 0;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANHRSKASMT 60
1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANHRSKASMT 60
DB 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPFNVSILVAMYST 120
61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPFNVSILVAMYST 120
QY 121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
QY 181 MONEAADATLVFVIGVVPALATLVLSRVRREDTPLDRDTGRLEPSAHLVAVTC 240
181 MONEAADATLVFVIGVVPALATLVLSRVRREDTPLDRDTGRLEPSAHLVAVTC 240
DB 241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
QY 241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
DB 301 QSFPSKLGRLMKKLPFGDRHCSPDHMGVQOYLA 333
301 QSFPSKLGRLMKKLPFGDRHCSPDHMGVQOYLA 333

RESULT 10
ABR56304
ID ABR56304 standard; protein; 714 AA.
XX
AC ABR56304;
XX
DT 20-NOV-2003 (first entry)
XX
DE pc90IHISGalpha6 protein.
XX
KW Human; anorectic; antidiabetic; antilipemic; hypothalamus;
G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
cibophobia; anorexia nervosa.
XX

OS Unidentified.

XX W02003030936-A1.
FN 17-APR-2003.
PD 02-OCT-2002; 2002WO-JP010250.
PF 02-OCT-2001; 2001JP-00306872.
PR 02-OCT-2001; 2001JP-00306872.
XX (SUMU) SUMITOMO PHARM CO LTD.
PA Suguru E, Tsuchida A, Yamana M, Taiji M;
PI WPI; 2003-354886/33.
DR N-PSDB; ACC70859.
XX
PT Inhibitors of expression or activity of G-protein coupled receptor 901
for treatment of lifestyle-related diseases and cibophobia.
XX
PS Disclosure; Page 73-75; 91pp; Japanese.
XX The present invention relates to novel remedies for the treatment of
diseases containing as an active component an inhibitor of the expression
or activity of hypothalamus-expressed G-protein coupled receptor 901 and
CC for treatment of cibophobia containing as an active component a
CC potentiator of the expression or activity of G-protein coupled receptor
901. The diseases which can be treated include obesity, diabetes and
CC hyperlipemia, and cibophobia (anorexia nervosa). The present sequence
was used to illustrate the invention
XX
SQ Sequence 714 AA;

Query Match 100.0%; Score 1733; DB 6; Length 714;
Best Local Similarity 100.0%; Pred. No. 1.1e-176; Indels 0; Gaps 0;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANHRSKASMT 60
1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANHRSKASMT 60
DB 61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPFNVSILVAMYST 120
61 MPDYVFVNMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVVALQIPFNVSILVAMYST 120
QY 121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
121 ALSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
QY 181 MONEAADATLVFVIGVVPALATLVLSRVRREDTPLDRDTGRLEPSAHLVAVTC 240
181 MONEAADATLVFVIGVVPALATLVLSRVRREDTPLDRDTGRLEPSAHLVAVTC 240
DB 241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
QY 241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
241 TORGLWTPHYLLILGHTVVISRGPVDANYLGILHFKVDPKSLAFSSSVTPPLLYRYNM 300
DB 301 QSFPSKLGRLMKKLPFGDRHCSPDHMGVQOYLA 333
301 QSFPSKLGRLMKKLPFGDRHCSPDHMGVQOYLA 333

RESULT 11
ABR56306
ID ABR56306 standard; protein; 735 AA.
XX
AC ABR56306;
XX
DT 20-NOV-2003 (first entry)
XX
DE pc90IHISGalpha2 protein.
XX
KW Human; anorectic; antidiabetic; antilipemic; hypothalamus;

KW G-protein coupled receptor 901; obesity; diabetes; hyperlipaemia;
XX cibophobia; anorexia nervosa.
OS Unidentified.
XX WO2003030936-A1.
XX PD 17-APR-2003.
XX PF 02-OCT-2002; 2002WO-JP010250.
XX PR 02-OCT-2001; 2001JP-00306872.
XX PA (SUMU) SUMITOMO PHARM CO LTD.
XX PI Suguru E, Teuchida A, Yamanaoka M, Taiji M;
XX DR WPI; 2003-354886/33.
XX DR N-PSDB; ACC70861.
XX PT Inhibitors of expression or activity of G-protein coupled receptor 901
XX PT for treatment of lifestyle-related diseases and cibophobia.
XX PS Disclosure; Page 85-87; 91pp; Japanese.
XX CC The present invention relates to novel remedies for the treatment of
XX CC diseases containing an active component an inhibitor of the expression
XX CC or activity of hypothalamus-expressed G-protein coupled receptor 901 and
XX CC for treatment of cibophobia containing as an active component a
XX CC potentiator of the expression or activity of G-protein coupled receptor
XX CC 901. The diseases which can be treated include obesity, diabetes and
XX CC hyperlipaemia, and cibophobia (anorexia nervosa). The present sequence
XX CC was used to illustrate the invention
SQ Sequence 735 AA;
Query Match 100.0%; Score 1733; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.2e-176;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MMSCSWENGTLGVEELPACODLQIGLSLILGLVGVGVGVCYNALLVLANLHRSKASMT 60
DB 1 MMSCSWENGTLGVEELPACODLQIGLSLILGLVGVGVGVCYNALLVLANLHRSKASMT 60
QY 61 MPDYVFVNNAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNNAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTTRAVCGFWGALITSPSSILFYICSHVSTRALBCKA 180
DB 121 ALLSLDHYIERALPRTYMASVYNTTRAVCGFWGALITSPSSILFYICSHVSTRALBCKA 180
QY 181 MONAERADATLVFIVGVVPALATLVALLSVRRREDTFLDRTGRLEPSARLLVAATVC 240
DB 181 MONAERADATLVFIVGVVPALATLVALLSVRRREDTFLDRTGRLEPSARLLVAATVC 240
QY 241 TQFGMTPHYLILGLHTTVIISRGKPVDAHYIGLHFVDFSKLAFSSFTPLLYRYMN 300
DB 241 TQFGMTPHYLILGLHTTVIISRGKPVDAHYIGLHFVDFSKLAFSSFTPLLYRYMN 300
QY 301 QSFPSKQLRLMKKLPCCGDRHCGSPDHMGVOQVLA 333
DB 301 QSFPSKQLRLMKKLPCCGDRHCGSPDHMGVOQVLA 333

RESULT 12
AA57285
XX AA57285 standard; protein; 333 AA.
AC AA57285;
XX
DT 05-JUN-2000 (first entry)
XX

DE Human GPCR protein (HGPRP) sequence (clone ID 2214673).
XX
KW Human; G protein coupled protein receptor; HGPRP; cell proliferation;
KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;
KW anti-atherosclerotic; hepatotropic; antiinflammatory; virocidic; leukemia;
KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.
OS Homo sapiens.
XX
PN WO200015793-A2.
XX PD 23-MAR-2000.
XX PF 17-SEP-1999; 99WO-US020958.
XX PR 17-SEP-1998; 98US-00156513.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;
XX PI Baughn MR;
XX DR WPI; 2000-271432/23.
XX DR N-PSDB; AA290523.
XX PT Human G protein coupled protein receptor peptides useful for the
XX PT prevention, diagnosis and treatment of cell proliferative, neurological
XX PT and immune disorders.
XX PS Claim 1; Page 68-69; 71pp; English.
XX
XX CC The invention provides human G protein coupled protein receptor (HGPRP)
XX CC polypeptides and polynucleotides encoding them. The polypeptides can be
XX CC produced by standard recombinant methodology. The polynucleotides and
XX CC polypeptides may be used in the prevention, treatment and diagnosis of
XX CC diseases associated with their inappropriate expression. Diseases that
XX CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
XX CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
XX CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
XX CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
XX CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as
XX CC diagnostic agents for detecting the presence of HGPRP polypeptides in
XX CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
SQ Sequence 333 AA;
Query Match 99.6%; Score 1726; DB 3; Length 333;
Best Local Similarity 99.7%; Pred. No. 2.2e-176;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MMSCSWENGTLGVEELPACODLQIGLSLILGLVGVGVGVCYNALLVLANLHRSKASMT 60
DB 1 MMSCSWENGTLGVEELPACODLQIGLSLILGLVGVGVGVCYNALLVLANLHRSKASMT 60
QY 61 MPDYVFVNNAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNNAVAGLVLSALAPVHLLGPSSRNALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTTRAVCGFWGALITSPSSILFYICSHVSTRALBCKA 180
DB 121 ALLSLDHYIERALPRTYMASVYNTTRAVCGFWGALITSPSSILFYICSHVSTRALBCKA 180
QY 181 MONAERADATLVFIVGVVPALATLVALLSVRRREDTFLDRTGRLEPSARLLVAATVC 240
DB 181 MONAERADATLVFIVGVVPALATLVALLSVRRREDTFLDRTGRLEPSARLLVAATVC 240
QY 241 TQFGMTPHYLILGLHTTVIISRGKPVDAHYIGLHFVDFSKLAFSSFTPLLYRYMN 300
DB 241 TQFGMTPHYLILGLHTTVIISRGKPVDAHYIGLHFVDFSKLAFSSFTPLLYRYMN 300
QY 301 QSFPSKQLRLMKKLPCCGDRHCGSPDHMGVOQVLA 333

DB 301 QSFPSKLGRLMKKLPCCGDRHCSPDHMGVQVLA 333

RESULT 13

AAU25584

ID AAU25584 standard; protein; 333 AA.

AAU25584;

18-DEC-2001 (first entry)

Human G Protein-Coupled Receptor (GPCR) polypeptide #31.

Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia; attention deficit disorder; anxiety; depression; bipolar disorder; neurological disorder; Huntington's disease; dementia; obesity; anorexia; metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis; type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer; cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV; viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer; antidepressant; anorectic; gene therapy.

Homo sapiens.

MO200162797-A2.

30-AUG-2001.

23-FEB-2001; 2001MO-US005676.

23-FEB-2000; 2000US-0184347P.

23-FEB-2000; 2000US-0184303P.

23-FEB-2000; 2000US-0184304P.

23-FEB-2000; 2000US-0184305P.

23-FEB-2000; 2000US-0184397P.

23-FEB-2000; 2000US-0186457P.

23-MAR-2000; 2000US-0186810P.

09-MAR-2000; 2000US-0188064P.

13-MAR-2000; 2000US-0188880P.

23-APR-2000; 2000US-0194344P.

23-JUN-2000; 2000US-0213616P.

11-JUL-2000; 2000US-0217369P.

11-JUL-2000; 2000US-0217370P.

14-JUL-2000; 2000US-0218337P.

20-JUL-2000; 2000US-0218492P.

(PHAA) PHARMACIA & UPJOHN CO.

Vogel G, Wood LS, Parodi LA, Lind P;

WPI; 2001-570628/64.

N-PSDB; AAS42836.

New isolated nucleic acid encoding a new G-protein coupled receptor

polypeptide for detecting receptor modulators that can treat mental

disorders, such as schizophrenia, anxiety, depression, or obesity.

Claim 35; Page 81; 279pp; English.

Sequences AAU25554-AAU25616 represent human G-protein coupled receptor

(GPCR) polypeptides of the invention. The proteins and their associated

DNA sequences can be used to identify compounds which bind to GPCR

polypeptides and in screening for compounds that modulate GPCR activity.

By screening a human subject for the presence of mutations in GPCR DNA, a

GPCR-related disorder or a genetic predisposition can be diagnosed. The

sequences can also be used for treatment and prevention of mental

disorders such as schizophrenia, attention deficit disorder, anxiety,

depression, dementia and bipolar disorder, neurological disorders such as

Huntington's disease, Parkinson's disease and Tourette's syndrome,

metabolic disorders such as obesity, anorexia and type 2 diabetes,

cardiovascular disorders such as thrombosis, myocardial infarction,

cardiomyopathy and atherosclerosis, viral infections caused by HIV and

CC cancers
XX
SQ Sequence 333 AA;

Query Match 99.6%; Score 1726; DB 4; Length 333;
Best Local Similarity 99.7%; Pred. No. 2.2e-176;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

1 MMSGWFNGTGVLEELPAQODLQGLSLSLGLVGVGVGVCYNALLVLANHSHKASMT 60

XX The invention relates to an isolated cDNA encoding G-protein coupled
CC receptor (GPCR) appearing as ABUS4865-ABUS4870. Also included are
CC fragments of the cDNAs, species variants having at least 75% identity to
CC the cDNAs, vectors comprising the cDNAs, a host cells comprising the
CC above vectors, producing a protein (comprising the protein), using a
CC cell under expression conditions and recovering the protein, using a
CC cDNA to detect expression of a nucleic acid in a sample or to screen for
CC compounds or molecules which bind to the cDNAs, using the GPCR proteins
CC to screen compounds or molecules for ligands, using a GPCR protein to
CC prepare and purify antibodies, an anti-GPCR antibody and using the
CC antibody to detect expression of a GPCR protein in a sample and its
CC diagnostic of cancer. The invention is useful for the diagnosis,
CC prognosis, treatment and evaluation of therapies for neoplastic,
CC neurological and immune disorders, particularly follicular carcinoma of
CC the thyroid, leiomyoma of the uterus, pancreatic cancer, epilepsy,
CC interstitial nephritis and immune response as a complication of cancer.
XX The present sequence is a human GPCR protein of the invention

Sequence 333 AA;

Query Match 99.6%; Score 1726; DB 5; Length 333;
Best Local Similarity 99.7%; Pred. No. 2.2e-176;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWSGFWNGTGLVEELPAACODLQGLSLSLGLVGVPGVLCYNALIVLANLHRSKSM 60
DB 1 MWSGFWNGTGLVEELPAACODLQGLSLSLGLVGVPGVLCYNALIVLANLHRSKSM 60
QY 61 MPDYFVNMVAVAGVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYFVNMVAVAGVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120
QY 121 ALLSLDHYIERALPRTWASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALAECAK 180
DB 121 ALLSLDHYIERALPRTWASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALAECAK 180
QY 181 MONNEADATLVFVIGYVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONNEADATLVFVIGYVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TOFGLMTPHYLILGHTVILISRGKPDVAHYLGILHFKVDSKLLAFSSSFVPLLYRYNN 300
DB 241 TOFGLMTPHYLILGHTVILISRGKPDVAHYLGILHFKVDSKLLAFSSSFVPLLYRYNN 300
QY 301 QSPSSKLGRLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSPSSKLGRLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 15

AA78856
ID AAY78856 standard; protein; 333 AA.

XX AAY78856;

DT 19-MAY-2000 (first entry)

XX Human chemokine receptor amino acid sequence.

XX Chemokine receptor; interleukin-8 compound inhibitor; chromosome 7p22;
KW inflammation; wound healing; neutropenia; myeloid leukemia; tumour;
KM toxin delivery; hypermegakaryocytopoietic disease; polycythaemia vera.
XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 11 /label= Unknown

FT Misc-difference 17 /note= "Encoded by RGG"

FT /label= Unknown
FT /note= "Encoded by MCT"

XX W0200000515-A2.
XX 06-JAN-2000.
PD 29-JUN-1999; 99MO-US012829.
PF 29-JUN-1999; 99MO-US012829.
XX 29-JUN-1999; 98US-00106800.
PR 22-JAN-1999; 99US-00236166.
XX (HYSE-) HYSEQ INC.
XX NPI; 2000-170907/15.
DR N-PSDB; AAZ90175.
XX New nucleic acid encoding chemokine receptor, useful for diagnosis and
PT treatment of e.g. neutropenia, inflammation and leukemia.
XX Claim 9; Fig 2; 138pp; English.

This sequence represents a human chemokine receptor amino acid sequence.
CC The nucleotide sequence encoding this protein is derived from a human
CC foetal liver-spleen cDNA library. The chemokine receptor inhibits the
CC activity of interleukin-8-type compounds through competition for cell
CC binding sites. The chemokine receptor gene is located on the short arm of
CC chromosome 7 at 7p22. The polynucleotide encoding the chemokine receptor
CC is useful as a hybridization probe or a PCR primer, the nucleotide
CC sequence may also be used for chromosome/gene mapping or in the
CC recombinant production of polypeptides and the production of antisense or
CC triple-forming molecules for the control of gene expression. The
CC chemokine receptor polypeptides are used to raise specific antibodies,
CC also for purification, detection or modulation of interleukin-8-type
CC chemokines (for diagnosis or prognosis, or monitoring chemokine
CC recruitment at a site of infection or inflammation). The protein sequence
CC can also be used as molecular weight markers or food supplements, and to
CC screen compound libraries for specific binding agents, potential agonists
CC or antagonists. Antibodies raised against the chemokine receptor
CC polypeptide sequence are used to detect or purify the polypeptide, also
CC for the diagnosis and treatment of activated or inflamed cells or
CC tissues, and to promote the healing of wounds. The polypeptide and
CC antibodies are also used to prevent neutropenia (associated with
CC chemotherapy or radiation treatment to protect myeloid precursors),
CC inflammation or other immune responses; also conditions associated with
CC hyperproliferation of progenitor cells (e.g. some myelogenous leukemias,
CC polycythaemia vera and hypermegakaryocytopoietic diseases). The
CC antibodies are potentially useful therapeutically, e.g. to carry toxins
CC to tumour cells

Sequence 333 AA;

Query Match 99.1%; Score 1717; DB 3; Length 333;
Best Local Similarity 99.4%; Pred. No. 2.1e-175;
Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MWSGFWNGTGLVEELPAACODLQGLSLSLGLVGVPGVLCYNALIVLANLHRSKSM 60
DB 1 MWSGFWNGTGLVEELPAACODLQGLSLSLGLVGVPGVLCYNALIVLANLHRSKSM 60
QY 61 MPDYFVNMVAVAGVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYFVNMVAVAGVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFNVSILVAMYST 120
QY 121 ALLSLDHYIERALPRTWASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALAECAK 180
DB 121 ALLSLDHYIERALPRTWASVYNTRHVCGFWGALLTSFSSILFYICSHVSTRALAECAK 180
QY 181 MONNEADATLVFVIGYVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONNEADATLVFVIGYVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TOFGLMTPHYLILGHTVILISRGKPDVAHYLGILHFKVDSKLLAFSSSFVPLLYRYNN 300
DB 241 TOFGLMTPHYLILGHTVILISRGKPDVAHYLGILHFKVDSKLLAFSSSFVPLLYRYNN 300

Qy 301 QSPSKLQRLMKLPCGDRHCSPPDHMGVQOVL 333
Db 301 QSPSKLQRLMKLPCGDRHCSPPDHMGVQOVL 333

Search completed: December 14, 2004, 20:23:05
Job time : 56.3965 secg

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:25:14 ; Search time 15.535 Seconds
(without alignments)
1421.692 Million cell updates/sec

Title: US-09-995-225B-2
Perfect score: 1733
Sequence: 1 MMSCSWNGTGLVEBELPACQ.....LPCGDHRCSPDHMGVQVLA 333

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCtus.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1711	98.7	333	4	US-09-409-778-3 Sequence 3, Appl1
2	271	15.6	375	2	US-08-480-994-38 Sequence 38, Appl1
3	271	15.6	375	2	US-08-616-844-38 Sequence 38, Appl1
4	271	15.6	375	2	US-08-559-654-38 Sequence 38, Appl1
5	271	15.6	375	2	US-08-485-573-38 Sequence 38, Appl1
6	271	15.6	375	3	US-08-944-868A-38 Sequence 38, Appl1
7	271	15.6	375	3	US-08-944-423A-38 Sequence 38, Appl1
8	271	15.6	375	3	US-08-925-743-38 Sequence 38, Appl1
9	271	15.6	375	3	US-08-944-496-38 Sequence 38, Appl1
10	271	15.6	375	3	US-08-925-767-38 Sequence 38, Appl1
11	269	15.5	375	4	US-09-170-496D-194 Sequence 194, App
12	269	15.5	375	4	US-09-170-496D-202 Sequence 202, App
13	267	15.4	375	4	US-09-711-889-2 Sequence 2, Appl1
14	267	15.4	375	4	US-09-170-068-2 Sequence 2, Appl1
15	267	15.4	375	4	US-09-170-068-2 Sequence 56, Appl1
16	267	15.4	375	4	US-09-170-496D-56 Sequence 70, Appl1
17	200	11.5	355	1	US-07-759-568-1 Sequence 1, Appl1
18	200	11.5	355	1	US-08-450-393A-8 Sequence 8, Appl1
19	200	11.5	355	2	US-08-390-000A-5 Sequence 5, Appl1
20	200	11.5	355	3	US-08-446-669-8 Sequence 8, Appl1
21	200	11.5	355	5	US-09-625-573-8 Sequence 8, Appl1
22	200	11.5	355	5	PCT-US95-00476-8 Sequence 8, Appl1
23	200	11.5	360	1	US-08-202-056-7 Sequence 8, Appl1
24	200	11.5	360	1	US-09-409-778-4 Sequence 4, Appl1
25	196.5	11.3	312	5	US-08-118-270-38 Sequence 38, Appl1
26	196.5	11.3	312	5	PCT-US93-08528-38 Sequence 38, Appl1
27	193.5	11.2	350	1	US-08-202-056-1 Sequence 1, Appl1

28	193.5	11.2	350	1	US-08-076-093A-2 Sequence 2, Appl1
29	193.5	11.2	350	1	US-08-450-393A-7 Sequence 7, Appl1
30	193.5	11.2	350	1	US-08-410-453A-1 Sequence 1, Appl1
31	193.5	11.2	350	1	US-08-701-265-2 Sequence 2, Appl1
32	193.5	11.2	350	1	US-08-410-454A-1 Sequence 1, Appl1
33	193.5	11.2	350	2	US-08-284-586-2 Sequence 2, Appl1
34	193.5	11.2	350	2	US-08-410-456A-1 Sequence 1, Appl1
35	193.5	11.2	350	2	US-08-805-478-2 Sequence 2, Appl1
36	193.5	11.2	350	2	US-08-802-627A-2 Sequence 2, Appl1
37	193.5	11.2	350	2	US-08-801-238-2 Sequence 2, Appl1
38	193.5	11.2	350	2	US-08-801-228-2 Sequence 2, Appl1
39	193.5	11.2	350	3	US-09-104-296-2 Sequence 2, Appl1
40	193.5	11.2	350	3	US-08-446-669-7 Sequence 7, Appl1
41	193.5	11.2	350	4	US-09-625-573-7 Sequence 7, Appl1
42	193.5	11.2	350	5	PCT-US95-00476-7 Sequence 7, Appl1
43	191.5	11.1	365	4	US-09-503-219B-8 Sequence 8, Appl1
44	188.5	10.9	365	4	US-10-039-659A-10 Sequence 10, Appl1
45	184.5	10.6	374	3	US-09-045-583-48 Sequence 48, Appl1

ALIGNMENTS

RESULT 1
US-09-409-778-3

Sequence 3, Application US/09409778

Patent No. 6472173

GENERAL INFORMATION:

APPLICANT: Ford, John

APPLICANT: Yeung, George

TITLE OF INVENTION: A NOVEL CHEMOKINE RECEPTOR OBTAINED FROM

FILE REFERENCE: 20411-742CON2 (now 28110/36057B)

CURRENT FILING DATE: 1999-09-22

PRIOR APPLICATION NUMBER: US/09/409,778

PRIOR FILING DATE: 1999-06-29

PRIOR APPLICATION NUMBER: US 09/236,166

PRIOR FILING DATE: 1999-01-22

PRIOR APPLICATION NUMBER: US 09/106,800

PRIOR FILING DATE: 1998-06-26

NUMBER OF SEQ ID NOS: 25

SOFTWARE: FaastSeq for Windows Version 3.0

SEQ ID NO 3

LENGTH: 333

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(333)

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-409-778-3

Query Match 98.7% Score 1711; DB 4; Length 333;

Best Local Similarity 99.1% Pred. No. 7e-150; Index 0; Gaps 0;

Matches 330; Conservative 0; Mismatches 3; Indels 0;

QY	1	MMSCSWNGTGLVEBELPACQDLOGLSLSLGLVGVPGVLCYNALLVLANLHRSKASMT	60
DB	1	MMSCSWNGTGLVEBELPACQDLOGLSLSLGLVGVPGVLCYNALLVLANLHRSKASMT	60
QY	61	MPDVYFYNNMAVAGLVTSALAPVHLGPPSSRRMLMSVGEVHALQIPFNVSLLVMYST	120
DB	61	MPDVYFYNNMAVAGLVTSALAPVHLGPPSSRRMLMSVGEVHALQIPFNVSLLVMYST	120
QY	121	ALLSLDHYIERALPRTYMASVYTRHYCGVWGAGALLTSFSSLLFTYICSHVSTRALCAK	180
DB	121	ALLSLDHYIERALPRTYMASVYTRHYCGVWGAGALLTSFSSLLFTYICSHVSTRALCAK	180
QY	181	MONAENADATLVIGVVPALATLVYLSVRREDTPLDRDTGRLEPSAHLVAATVC	240
DB	181	MONAENADATLVIGVVPALATLVYLSVRREDTPLDRDTGRLEPSAHLVAATVC	240

QY 241 TOGFWTPHYLLIGHTVVISRSGKPYDAHYLGLHFVQKPSKLLAFSSSVTEPLRYXNN 300
DB 241 TOGFWTPHYLLIGHTVVISRSGKPYDAHYLGLHFVQKPSKLLAFSSSVTEPLRYXNN 300

QY 301 QSFPSKLGRLMKKLPCCGRHCSPDHMGVQVLA 333
DB 301 QSFPSKLGRLMKKLPCCGRHCSPDHMGVQVLA 333

RESULT 2

US-08-480-994-38
; Sequence 38, Application US/08480994
; Patent No. 5834248

; GENERAL INFORMATION:

; APPLICANT: FALB, DEAN A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 Avenue of the Americas
; CITY: New York

; STATE: New York
; COUNTRY: USA

; ZIP: 10036-2711

; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,994

; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573

; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995

; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-033

; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864

; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 38:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids

; TYPE: amino acid
; STRANDEDNESS:

; TOPOLOGY: unknown
; MOLECULE TYPE: protein

US-08-480-994-38

Query Match 15.6%; Score 271; DB 2; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;

Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACODLQGLSLISLGLVGVGVCGLCNALVLANLSKASMTMPDVYV 67
DB 44 NGTG--ELSEHQOYVIGL-FLSCLYITFLFPIGVGNILIVNLSFEEKMTIDPLYFI 99

QY 68 NNAVAGLVLSALAPVHLGLPPSSRWALMSVGGEVH-----VALQIFP-NVSLVAMYSTRA 121
DB 100 NNAVADLLIVADSLIEVF-----NLHRRYDIAVLCTFMSLFLAVNNYSSV 145

QY 122 L-----LSLDHYIRRALPRITMASVYNTRH---VCGFWMGALLTSFSSLSLFTICSHVST 173
DB 146 FFLTWMSFDRIYI--ALARAARCSLFRTKHARLSGGLIW---NASVSATLVPTAVHLQH 200

QY 174 RALECARKMONAEEADATVIFIGVYV--ALATLYVALILSRVREDTPIURDTRGLEP-- 229
DB 201 TBEACCFPADVVEVQMLEVTLGFIYFPAIIGCYLIVAVYVRAH-----RRRG-LRPRR 254

QY 230 -SAHRLVAVTCTQPGFWTPHYLLIGHTVVISRGP-----VDHAYLGLHFVKD 279
DB 255 QALARMILAVLVVFFVCWLPENVFISVH--LLQRPQGAAPCKQSFRRHAPL-TGHIV-- 309

QY 280 FSKLLAFSSSVTEPLRYXNNOSFPSKLGRLMKK---LPCGRHCSPDHMGVQVLA 332
DB 310 -NLAFNSNCINPLIYSFLGETFRDKLRLYEOKTNLPALDRFC--HAALKAVI 360

RESULT 3

US-08-616-844-38

; Sequence 38, Application US/08616844
; Patent No. 5849578

; GENERAL INFORMATION:

; APPLICANT: FALB, DEAN A.

; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
; TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 Avenue of the Americas
; CITY: New York

; STATE: New York
; COUNTRY: USA

; ZIP: 10036-2711

; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,844

; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654

; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A.

; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-053

; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864

; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 38:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids

; TYPE: amino acid
; STRANDEDNESS:

; TOPOLOGY: unknown
; MOLECULE TYPE: protein

US-08-616-844-38

Query Match 15.6%; Score 271; DB 2; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;

Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACODLQGLSLISLGLVGVGVCGLCNALVLANLSKASMTMPDVYV 67

DB 44 NGTG--ELSEHQOYVIGL-FLSCLYITFLFPIGVGNILIVNLSFEEKMTIDPLYFI 99

QY 66 NNAVAGLVLSALAPVHLGLPPSSRWALMSVGGEVH-----VALQIFP-NVSLVAMYSTRA 121

Db 100 NLAVADILVADSLIEVF-----NHERYDIAVLCFMSLFELRNWYSSV 145
Qy 122 L-----LSLDHYIERALPRTMASVYNTRH-----VCGFWGALLTSSSLFYCSHVT 173
Db 146 FLITWMSFDRIYI--ALAPARACSLFRTKHARLSGGLIM--MASVATLVPTAVHLOH 200
Qy 174 RALECAGMNAEADATLVFIQVVP--ALATYALVLSRVREDPTLDRDGRLEP-- 229
Db 201 TDEACFCFADVRERVOMLEVTIGFIVPFAIIGLCYSLIVRVLVRAH-----RHRC-LRPFR 254
Qy 230 -SAHRLIVATVCTQFGIMTPHYILLGHTVYISRGK-----VDHYLGILHFVKD 279
Db 255 QKALRMILAVLVFVFCWLPENVFISVH--LQRTQGAAPCKQSPFHAPHL-TGHTV-- 309
Qy 280 FSKLAFSSSFVTPLLYRYMNSFPKQLQRLMK--LPCGDRHCSPDHMGVOVL 332
Db 310 --NLAAFSNSCLNPLIYFSGTFRDRLRYIEQKTNLPALDRFC--HAAKXAVI 360

RESULT 4
US-08-599-654-38
Sequence 38, Application US/08599654
Patent No. 5882925
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
NUMBER OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
CORRESPONDENCE ADDRESSES: 54
ADDRESS: PENNIE & EDWARDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,654
FILING DATE: 09-FEB-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-041
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-599-654-38

Query Match 15.6%; Score 271; DB 2; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

Qy 8 NGTGLVELPACDQLGLSLSLGLVGVPGVGLCNALLVLANHKSASMTMPDYFYV 67

Db 44 NGTG---ELSEHQYVIGL-FLSCLYTIFFEPFGFNILLVNVSFRKMTIPDLXYI 99
Qy 68 NNAVAGLIVSALAPVHILGPPSRMALMSVGVH-----VAQIIP-NVSSLYANSTA 121
Db 100 NLAVADILVADSLIEVF-----NHERYDIAVLCFMSLFELRNWYSSV 145
Qy 122 L-----LSLDHYIERALPRTMASVYNTRH-----VCGFWGALLTSSSLFYCSHVT 173
Db 146 FLITWMSFDRIYI--ALAPARACSLFRTKHARLSGGLIM--MASVATLVPTAVHLOH 200
Qy 174 RALECAGMNAEADATLVFIQVVP--ALATYALVLSRVREDPTLDRDGRLEP-- 229
Db 201 TDEACFCFADVRERVOMLEVTIGFIVPFAIIGLCYSLIVRVLVRAH-----RHRC-LRPFR 254
Qy 230 -SAHRLIVATVCTQFGIMTPHYILLGHTVYISRGK-----VDHYLGILHFVKD 279
Db 255 QKALRMILAVLVFVFCWLPENVFISVH--LQRTQGAAPCKQSPFHAPHL-TGHTV-- 309
Qy 280 FSKLAFSSSFVTPLLYRYMNSFPKQLQRLMK--LPCGDRHCSPDHMGVOVL 332
Db 310 --NLAAFSNSCLNPLIYFSGTFRDRLRYIEQKTNLPALDRFC--HAAKXAVI 360

RESULT 5
US-08-485-573-38
Sequence 38, Application US/08485573
Patent No. 5968770
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
NUMBER OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
CORRESPONDENCE ADDRESSES: 38
ADDRESS: PENNIE & EDWARDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,573
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-032
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-573-38

Query Match 15.6%; Score 271; DB 2; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-944-423A-38

Query Match 15.6%; Score 271; DB 3; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17; Indels 70; Gaps 19;
Matches 98; Conservative 64; Mismatches 124;

QY 8 NGTGLVEELPACODLQGLSLSLGLVGVPGLCYNALVLANLHASKAMTPDYYFV 67
DB 44 NGTG---ELSEHQYVIGL-FLSCLYTIPLFPIGPNILILVNIISFRKMTIPDLYFI 99
QY 68 NMAVAGVLSALAPVHLGPPSSRMALMSVGSEVH-----VALQIP-NVSSLVAMYSTA 121
DB 100 NLAVADLILVADSLIEVF-----NLSHRYDIAVLCTFMSFLRVNMYSSV 145
QY 122 L-----LSLHYIERALPRTMASVYNTRH-----VCGFWGALLTSFSLFYICSHYST 173
DB 146 FFLTWSFDRIYI--ALARMRCSLFKTKHARLSCGLIW---MASVSATLVPTAVHLQH 200
QY 174 RALECAMQNAEADATLVFVGVV--ALATYALVLSRVREDTPLDRDGRLEP-- 229
DB 201 TDEACFCFADVREVQMLEVTLGFIIVPAIIGLCYSLIVRVLAH-----RHNG-LRPR 254
QY 230 -SAHRLIVATVCTQFGIMTPHYLLILGHTVVISRGP-----VDAHYLGLHFKVD 279
DB 255 OKALRMILAVLVFVCMLEBNVFSVH--LQRTQGAAPCKQSFHAHPL-TGHIV-- 309
QY 280 FSKLLAPSSFVPLRYVNNQSPFSLQRLMK--LPCGDHCSPDHNGVQOVL 332
DB 310 --NLAAFSNCLNPLIYSLFGETFRDLRLYIEQKTNLPALDRFC--HAALKAVI 360

RESULT 8

US-08-925-743-38
Sequence 38, Application US/08925743
Patent No. 6034558
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/925,743
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,573
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-032
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-925-743-38

Query Match 15.6%; Score 271; DB 3; Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17; Indels 70; Gaps 19;
Matches 98; Conservative 64; Mismatches 124;

QY 8 NGTGLVEELPACODLQGLSLSLGLVGVPGLCYNALVLANLHASKAMTPDYYFV 67
DB 44 NGTG---ELSEHQYVIGL-FLSCLYTIPLFPIGPNILILVNIISFRKMTIPDLYFI 99
QY 68 NMAVAGVLSALAPVHLGPPSSRMALMSVGSEVH-----VALQIP-NVSSLVAMYSTA 121
DB 100 NLAVADLILVADSLIEVF-----NLSHRYDIAVLCTFMSFLRVNMYSSV 145
QY 122 L-----LSLHYIERALPRTMASVYNTRH-----VCGFWGALLTSFSLFYICSHYST 173
DB 146 FFLTWSFDRIYI--ALARMRCSLFKTKHARLSCGLIW---MASVSATLVPTAVHLQH 200
QY 174 RALECAMQNAEADATLVFVGVV--ALATYALVLSRVREDTPLDRDGRLEP-- 229
DB 201 TDEACFCFADVREVQMLEVTLGFIIVPAIIGLCYSLIVRVLAH-----RHNG-LRPR 254
QY 230 -SAHRLIVATVCTQFGIMTPHYLLILGHTVVISRGP-----VDAHYLGLHFKVD 279
DB 255 OKALRMILAVLVFVCMLEBNVFSVH--LQRTQGAAPCKQSFHAHPL-TGHIV-- 309
QY 280 FSKLLAPSSFVPLRYVNNQSPFSLQRLMK--LPCGDHCSPDHNGVQOVL 332
DB 310 --NLAAFSNCLNPLIYSLFGETFRDLRLYIEQKTNLPALDRFC--HAALKAVI 360

RESULT 9

US-08-944-496-38
Sequence 38, Application US/08944496
Patent No. 6124433
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,496
FILING DATE: 06-OCT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A

REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELE: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-944-496-38

Query Match 15.6%; Score 271, DB 3, Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

8 NGTGLVEELPACODLQIGLSLILGLVGVPGVGLCNALLVLANLHASKASTMDDVYFV 67
44 NGTG---ELSEHQYVIGL-FLSCLYTFIFLPIGVGNILILVWNISFRKMTIPDLVFI 99
68 NNAVAGLVLSALAPVHLGPPSSRWALMSVGEVH-----VALQIPF-NVSSLVAMYSTA 121
100 NLAVALDLIVADSLIEVF-----NLHERRYDIAVLCTFMSLFLRVNMYSSV 145
QY 122 L-----LSLDHYIRALPRTYMASVYNTRH---VCGFWGALLTSFSSLIIFYCSHVST 173
DB 146 FFLTWMSFDRIY--ALARAMRCSLFRTHARLSGLIW---MASVSATLVPEFAVHLQH 200
QY 174 RALECAKQNAEAAADATVFIGYVVP--ALATYALVILSRVREDTPLDDRTGRLEP-- 229
DB 201 TDEACFCPADRVQWMLVETLGIYFPAIIGLCYSILRVLYRAH-----RHRG-LRRRR 254
QY 230 -SAHRLLVATVCTQFGMLTPHYLLIGHTVIISRGK-----VDANYLGILHFPYKD 279
DB 255 QKALRMILAVLVLFVFCWLPENVFISVH--LLQRTQPGAPCKQSFRAHPL-TGHIV-- 309
QY 280 FSKLLAFSSSFVTPILRYNMQSPFSKIQRLMK---LPCGDRHSPHHMGVQOVL 332
DB 310 --NLAAFNSCLMPLIYSFLGETFRDKLRLYIEQKTNLPALDRFC--HAALKAVI 360

RESULT 10
US-08-925-767-38
Sequence 38, Application US/08925767
Patent No. 6225084
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSER: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/925,767
FILING DATE: 09-SEPT-1997
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELE: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-925-767-38

Query Match 15.6%; Score 271, DB 3, Length 375;
Best Local Similarity 27.5%; Pred. No. 6,7e-17;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

8 NGTGLVEELPACODLQIGLSLILGLVGVPGVGLCNALLVLANLHASKASTMDDVYFV 67
44 NGTG---ELSEHQYVIGL-FLSCLYTFIFLPIGVGNILILVWNISFRKMTIPDLVFI 99
68 NNAVAGLVLSALAPVHLGPPSSRWALMSVGEVH-----VALQIPF-NVSSLVAMYSTA 121
100 NLAVALDLIVADSLIEVF-----NLHERRYDIAVLCTFMSLFLRVNMYSSV 145
QY 122 L-----LSLDHYIRALPRTYMASVYNTRH---VCGFWGALLTSFSSLIIFYCSHVST 173
DB 146 FFLTWMSFDRIY--ALARAMRCSLFRTHARLSGLIW---MASVSATLVPEFAVHLQH 200
QY 174 RALECAKQNAEAAADATVFIGYVVP--ALATYALVILSRVREDTPLDDRTGRLEP-- 229
DB 201 TDEACFCPADRVQWMLVETLGIYFPAIIGLCYSILRVLYRAH-----RHRG-LRRRR 254
QY 230 -SAHRLLVATVCTQFGMLTPHYLLIGHTVIISRGK-----VDANYLGILHFPYKD 279
DB 255 QKALRMILAVLVLFVFCWLPENVFISVH--LLQRTQPGAPCKQSFRAHPL-TGHIV-- 309
QY 280 FSKLLAFSSSFVTPILRYNMQSPFSKIQRLMK---LPCGDRHSPHHMGVQOVL 332
DB 310 --NLAAFNSCLMPLIYSFLGETFRDKLRLYIEQKTNLPALDRFC--HAALKAVI 360

RESULT 11
US-09-170-496D-194
Sequence 194, Application US/09170496D
Patent No. 6555339
GENERAL INFORMATION:
APPLICANT: Behan, Dominic P.
TITLE OF INVENTION: Receptors
FILE REFERENCE: AREN-0040
CURRENT APPLICATION NUMBER: US/09/170,496D
CURRENT FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 294
SOFTWARE: PatentIn version 3.1
SEQ ID NO 194
LENGTH: 375
TYPE: PRT
ORGANISM: Homo sapiens
US-09-170-496D-194

Query Match 15.5%; Score 269, DB 4, Length 375;
Best Local Similarity 27.4%; Pred. No. 1e-16;
Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

```
QY      8 NGTGLVEELPACQDLOGLSLSLGLVGVPGVGLCYNALLVLANLHASKASMTMPDYFV 67
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      44 NGTG---ELSEHQOYVIGL-FLSCLYTIPLFPFGVGNILLVNISFRKMTIPDLXYFI 99
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      68 NMAVAGLVLSALAPVHLLGPPSSRMALMSVGEVHVALQIPNVSSLVANYSTAL----L 123
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      100 NLAVALDILVADSLIEFNFHERRYDAVLCITFMSLFLQ-----VNNYSSVFFLTWM 151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      124 SLDHYIERALPRTYMASVYNTNRH---VCGFWGAGALLTSFSSLLFYICSHVSTRALECA 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      152 SFDRYI--ALARMRCSLFRTKHARLSCGLIW---MASVSATLVPTAVHLQHTDBACF 206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      180 KMQNAEADATLVFIGVVP--ALATLYALVLSRVREDTPLDRTGRLEP---SAHRL 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      207 CFADVREVQWLTVGLTGFIVPAIIGLCYSLIVRLVRAH-----RHRG-LRPRRQKAKRM 260
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      235 LVATVCTQFGMLTPHYIILGHTVIISRGK-----VDANYGLLHFVQDFSLTA 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      261 ILAVLVVFPVCMLEPNVFIIVH--LLQRTQGAAPCKQSPRHAHPL-TGHIV---NLAA 313
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      286 FSSSFVTPLLYRYMNSFPKQLQRLMK--LPCGDHCSPDHMGVOQVL 332
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      314 FNSNCINPLIYSFLETFRDKLRLYIQKTNLPALNRF--HAALKAVI 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-170-496D-202
; Sequence 202, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 202
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-202

Query Match      15.4%; Score 269; DB 4; Length 375;
Best Local Similarity 27.4%; Pred. No. 1e-16;
Matches 96; Conservative 65; Mismatches 111; Indels 58; Gaps 17;

QY      8 NGTGLVEELPACQDLOGLSLSLGLVGVPGVGLCYNALLVLANLHASKASMTMPDYFV 67
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      44 NGTG---ELSEHQOYVIGL-FLSCLYTIPLFPFGVGNILLVNISFRKMTIPDLXYFI 99
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      68 NMAVAGLVLSALAPVHLLGPPSSRMALMSVGEVHVALQIPNVSSLVANYSTAL----L 123
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      100 NLAVALDILVADSLIEFNFHERRYDAVLCITFMSLFLQ-----VNNYSSVFFLTWM 151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      124 SLDHYIERALPRTYMASVYNTNRH---VCGFWGAGALLTSFSSLLFYICSHVSTRALECA 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      152 SFDRYI--ALARMRCSLFRTKHARLSCGLIW---MASVSATLVPTAVHLQHTDBACF 206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      180 KMQNAEADATLVFIGVVP--ALATLYALVLSRVREDTPLDRTGRLEP---SAHRL 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      207 CFADVREVQWLTVGLTGFIVPAIIGLCYSLIVRLVRAH-----RHRG-LRPRRQKAKRM 260
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      235 LVATVCTQFGMLTPHYIILGHTVIISRGK-----VDANYGLLHFVQDFSLTA 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      261 ILAVLVVFPVCMLEPNVFIIVH--LLQRTQGAAPCKQSPRHAHPL-TGHIV---NLTA 313
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      286 FSSSFVTPLLYRYMNSFPKQLQRLMK--LPCGDHCSPDHMGVOQVL 332
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      314 FNSNCINPLIYSFLETFRDKLRLYIQKTNLPALNRF--HAALKAVI 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 13
US-09-711-889-2
; Sequence 2, Application US/09711889
; Patent No. 6468769
; GENERAL INFORMATION:
; APPLICANT: OMNAN, CHRISTER
; TITLE OF INVENTION: HEPTAHRELIX RECEPTOR AND ITS USE
; FILE REFERENCE: 076750002 SEQUENCE LISTING
; CURRENT APPLICATION NUMBER: US/09/711,889
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 09/170,068
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/063,494
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-711-889-2

Query Match      15.4%; Score 267; DB 4; Length 375;
Best Local Similarity 27.4%; Pred. No. 1.6e-16;
Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

QY      8 NGTGLVEELPACQDLOGLSLSLGLVGVPGVGLCYNALLVLANLHASKASMTMPDYFV 67
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      44 NGTG---ELSEHQOYVIGL-FLSCLYTIPLFPFGVGNILLVNISFRKMTIPDLXYFI 99
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      68 NMAVAGLVLSALAPVHLLGPPSSRMALMSVGEVHVALQIPNVSSLVANYSTAL----L 123
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      100 NLAVALDILVADSLIEFNFHERRYDAVLCITFMSLFLQ-----VNNYSSVFFLTWM 151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      124 SLDHYIERALPRTYMASVYNTNRH---VCGFWGAGALLTSFSSLLFYICSHVSTRALECA 179
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      152 SFDRYI--ALARMRCSLFRTKHARLSCGLIW---MASVSATLVPTAVHLQHTDBACF 206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      180 KMQNAEADATLVFIGVVP--ALATLYALVLSRVREDTPLDRTGRLEP---SAHRL 234
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      207 CFADVREVQWLTVGLTGFIVPAIIGLCYSLIVRLVRAH-----RHRG-LRPRRQKAKRM 260
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      235 LVATVCTQFGMLTPHYIILGHTVIISRGK-----VDANYGLLHFVQDFSLTA 285
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      261 ILAVLVVFPVCMLEPNVFIIVH--LLQRTQGAAPCKQSPRHAHPL-TGHIV---NLAA 313
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      286 FSSSFVTPLLYRYMNSFPKQLQRLMK--LPCGDHCSPDHMGVOQVL 332
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      314 FNSNCINPLIYSFLETFRDKLRLYIQKTNLPALNRF--HAALKAVI 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-170-068-2
; Sequence 2, Application US/09170068
; Patent No. 6489442
; GENERAL INFORMATION:
; APPLICANT: OMNAN, CHRISTER
; TITLE OF INVENTION: HEPTAHRELIX RECEPTOR AND ITS USE
; FILE REFERENCE: 07675,0002
; CURRENT APPLICATION NUMBER: US/09/170,068
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 60/063,494
; EARLIER FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-068-2
```

Query Match 15.4%; Score 267; DB 4; Length 375;
 Best Local Similarity 27.4%; Pred. No. 1.6e-16;

Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

```

QY      8 NGGGLVEELPACODLQGLSLSLGLVGVPGLCYNALVLVIANLHASKASTMPDYYFV 67
DB      44 NGTG---ELSEHQOYVIGL-FLSCLTYTFLPFGVGNILIVNISFREKMTIDLYFI 99
QY      68 NMAVAGLVLSALAPVHLGPPSSRWALMSVGEVHALQIPNNVSLVAMYSTAL----L 123
DB      100 NLAVADLLIVADSLIEVNLHERRYDIAVLCTFMSLFLO-----VNMYSVFFLTWM 151
QY      124 SLDHYTEALPPTYNASVYNTRH---VCGFWGALITSPSLLFYICSHVSTRALECA 179
DB      152 SFDRYI--ALARAMCSLFRTHARLSGGLIM---MASVSATLVPTAVHLQHTDEACF 206
QY      180 KQONAEADATLVFGVVP--ALATYALVLRSVRREDTPLDRDGTGLEP---SAHRL 234
DB      207 CPADVREVQMLEVTIGFVFPALIGLCYSLIVRLVRAH---RHGG-LRPRQKALRM 260
QY      235 LVATVCTQFGMLTPHYLLLGHTVLIISRGK-----VDAPHGLHFPVKDFSKLLA 285
DB      261 ILAVLVFVFCWLPENVFISVH--LQRTQGAAPCKQSFRAHPL--TGHIV---NLAA 313
QY      286 FSSSFVTPLLYRYMNSQSPSKLQRLMKK--LPCGDRHCSPDHMGVOQVL 332
DB      314 FSNSCNPLIYSFLGETFRDKRLRYIEQKTNLPALNRFCC--HAALKAVI 360

```

RESULT 15

US-09-170-496D-56

; Sequence 56, Application US/09170496D

; Patent No. 6555339

; GENERAL INFORMATION:

; APPLICANT: Behan, Dominic P.

; APPLICANT: Chalmers, Derek T.

; APPLICANT: Liaw, Chen W.

; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-

; FILE REFERENCE: AREN-0040

; CURRENT APPLICATION NUMBER: US/09/170,496D

; NUMBER OF SEQ ID NOS: 294

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 56

; LENGTH: 375

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-170-496D-56

Query Match 15.4%; Score 267; DB 4; Length 375;

Best Local Similarity 27.4%; Pred. No. 1.6e-16;

Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

```

QY      8 NGGGLVEELPACODLQGLSLSLGLVGVPGLCYNALVLVIANLHASKASTMPDYYFV 67
DB      44 NGTG---ELSEHQOYVIGL-FLSCLTYTFLPFGVGNILIVNISFREKMTIDLYFI 99
QY      68 NMAVAGLVLSALAPVHLGPPSSRWALMSVGEVHALQIPNNVSLVAMYSTAL----L 123
DB      100 NLAVADLLIVADSLIEVNLHERRYDIAVLCTFMSLFLO-----VNMYSVFFLTWM 151
QY      124 SLDHYTEALPPTYNASVYNTRH---VCGFWGALITSPSLLFYICSHVSTRALECA 179
DB      152 SFDRYI--ALARAMCSLFRTHARLSGGLIM---MASVSATLVPTAVHLQHTDEACF 206
QY      180 KQONAEADATLVFGVVP--ALATYALVLRSVRREDTPLDRDGTGLEP---SAHRL 234
DB      207 CPADVREVQMLEVTIGFVFPALIGLCYSLIVRLVRAH---RHGG-LRPRQKALRM 260
QY      235 LVATVCTQFGMLTPHYLLLGHTVLIISRGK-----VDAPHGLHFPVKDFSKLLA 285
DB      261 ILAVLVFVFCWLPENVFISVH--LQRTQGAAPCKQSFRAHPL--TGHIV---NLAA 313

```

```

QY      286 FSSSFVTPLLYRYMNSQSPSKLQRLMKK--LPCGDRHCSPDHMGVOQVL 332
DB      314 FSNSCNPLIYSFLGETFRDKRLRYIEQKTNLPALNRFCC--HAALKAVI 360

```

Search completed: December 14, 2004, 20:28:12
 Job time : 16.5335 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 14, 2004, 20:23:14 ; Search time 46.6006 Seconds

(without alignments)
2552.337 Million cell updates/sec

Title: US-09-995-225B-2

Perfect score: 1733
Sequence: 1 MMSCSFNGTGLVEELPACO.....LPCCDRHCRPDHMGVQVLA 333Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*

2: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*

3: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*

4: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*

5: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*

6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*

8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*

9: /cgn2_6/prodata/1/pubppa/US09A_PUBCOMB.pep:*

10: /cgn2_6/prodata/1/pubppa/US09B_PUBCOMB.pep:*

11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*

12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*

13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*

14: /cgn2_6/prodata/1/pubppa/US10B_PUBCOMB.pep:*

15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*

16: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*

17: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*

18: /cgn2_6/prodata/1/pubppa/US11_NEW_PUB.pep:*

19: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*

20: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733	100.0	333	9	US-09-995-225-2
2	1733	100.0	333	10	US-09-791-932-114
3	1733	100.0	333	10	US-09-990-940-4
4	1733	100.0	333	10	US-09-995-225-2
5	1733	100.0	333	14	US-10-165-844-13
6	1733	100.0	333	14	US-10-101-148-3
7	1733	100.0	333	14	US-10-292-798-1016
8	1733	100.0	333	15	US-10-296-115-1174
9	1733	100.0	333	14	US-10-017-161-11202
10	1726	99.6	333	9	US-09-895-686-3
11	1726	99.6	333	10	US-09-791-932-91
12	1717	99.1	333	14	US-10-225-567A-680
13	271	15.6	375	9	US-09-371-900-38

14	271	15.6	375	9	US-09-924-417-59	Sequence 59, Appl
15	271	15.6	375	9	US-09-970-820-38	Sequence 38, Appl
16	271	15.6	375	9	US-09-986-718-38	Sequence 38, Appl
17	271	15.6	375	14	US-10-186-950-38	Sequence 38, Appl
18	271	15.6	375	15	US-10-653-872-59	Sequence 59, Appl
19	269	15.5	375	14	US-10-251-385-194	Sequence 194, Appl
20	269	15.5	375	14	US-10-251-385-202	Sequence 202, Appl
21	267	15.4	375	14	US-10-251-385-56	Sequence 56, Appl
22	267	15.4	375	14	US-10-251-385-70	Sequence 70, Appl
23	267	15.4	375	14	US-10-225-567A-92	Sequence 92, Appl
24	210.5	12.1	355	14	US-10-237-563-37	Sequence 37, Appl
25	208.5	12.0	355	14	US-10-237-563-34	Sequence 34, Appl
26	207.5	12.0	355	14	US-10-237-563-36	Sequence 36, Appl
27	201.5	11.6	355	14	US-10-237-563-35	Sequence 35, Appl
28	201	11.6	355	14	US-10-237-563-29	Sequence 29, Appl
29	201	11.6	355	14	US-10-237-563-30	Sequence 30, Appl
30	200	11.5	354	14	US-10-411-284-11	Sequence 11, Appl
31	200	11.5	355	10	US-09-893-512-10	Sequence 10, Appl
32	200	11.5	355	14	US-10-237-563-27	Sequence 27, Appl
33	200	11.5	355	14	US-10-237-563-28	Sequence 28, Appl
34	200	11.5	355	14	US-10-237-563-31	Sequence 31, Appl
35	200	11.5	355	17	US-10-791-592-8	Sequence 8, Appl
36	200	11.5	355	17	US-10-791-166-8	Sequence 8, Appl
37	200	11.5	360	13	US-10-087-192-882	Sequence 882, Appl
38	200	11.5	360	14	US-10-099-007A-16	Sequence 16, Appl
39	200	11.5	360	14	US-10-225-567A-386	Sequence 386, Appl
40	200	11.5	360	14	US-10-239-423-77	Sequence 77, Appl
41	200	11.5	360	14	US-10-101-148-4	Sequence 4, Appl
42	200	11.5	399	15	US-10-334-143-7	Sequence 7, Appl
43	200	11.5	399	15	US-10-276-774-2394	Sequence 2394, Ap
44	193.5	11.2	350	9	US-09-782-980-83	Sequence 83, Appl
45	193.5	11.2	350	9	US-09-884-430-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-995-225-2
Sequence 2, Application US/09995225
Publication No. US20020193584A1
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huang T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Priddy, Cameron
TITLE OR INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
FILE REFERENCE: ARN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917

;; PRIOR FILING DATE: 2001-05-14
;; PRIOR APPLICATION NUMBER: 60/309,208
;; PRIOR FILING DATE: 2001-07-31
;; NUMBER OF SEQ ID NOS: 67
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 333
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-2

Query Match 100.0%; Score 1733; DB 9; Length 333;
Best Local Similarity 100.0%; Pred. No. 9,8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGTGLVEELPACODIQLGLSLSLGLVGVPGVGLCYNALLVLANTLHASKASMT 60
DB 1 MMSCSWFGTGLVEELPACODIQLGLSLSLGLVGVPGVGLCYNALLVLANTLHASKASMT 60
QY 61 MPDYVFNMAVAGVLSALAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYVFNMAVAGVLSALAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
QY 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSPSSILFYICSHVSTRALBCK 180
DB 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSPSSILFYICSHVSTRALBCK 180
QY 181 MONEAADAATLVFIVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHRLLVAATVC 240
DB 181 MONEAADAATLVFIVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHRLLVAATVC 240
QY 241 TQFGMTPHYLLILGHTVYIISRGKPYDAHYIGLHFVDFSKLAFSSSFVTPPLLYRYMN 300
DB 241 TQFGMTPHYLLILGHTVYIISRGKPYDAHYIGLHFVDFSKLAFSSSFVTPPLLYRYMN 300
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 2
US-09-791-932-114
;; Sequence 114, Application US/09791932
;; Publication No. US20030003451A1
;; GENERAL INFORMATION:
;; APPLICANT: Vogeli, Gabriel
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Hiebsch, Ronald R.
;; APPLICANT: Lind, Peter
;; APPLICANT: Kaytes, Paul S.
;; APPLICANT: Ruff, Valerie
;; APPLICANT: Huff, Rita M.
;; APPLICANT: Wood, Linda S.
;; TITLE OF INVENTION: No. US20030003451A1el G Protein-Coupled Receptors Cross-Referen
;; FILE REFERENCE: 00325 US1
;; CURRENT APPLICATION NUMBER: US/09/791,932
;; CURRENT FILING DATE: 2001-02-23
;; PRIOR APPLICATION NUMBER: 60/184,305
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,304
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,303
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,397
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/184,247
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/188,880
;; PRIOR FILING DATE: 2000-03-13
;; PRIOR APPLICATION NUMBER: 60/217,369
;; PRIOR FILING DATE: 2000-07-11

;; PRIOR APPLICATION NUMBER: 60/217,370
;; PRIOR FILING DATE: 2000-07-11
;; PRIOR APPLICATION NUMBER: 60/218,492
;; PRIOR FILING DATE: 2000-07-20
;; PRIOR APPLICATION NUMBER: 60/186,810
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/188,064
;; PRIOR FILING DATE: 2000-03-09
;; PRIOR APPLICATION NUMBER: 60/186,457
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: 60/213,861
;; PRIOR FILING DATE: 2000-06-23
;; PRIOR APPLICATION NUMBER: 60/194,344
;; PRIOR FILING DATE: 2000-04-03
;; PRIOR APPLICATION NUMBER: 60/218,337
;; PRIOR FILING DATE: 2000-07-14
;; NUMBER OF SEQ ID NOS: 184
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 114
;; LENGTH: 333
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-791-932-114

Query Match 100.0%; Score 1733; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 9,8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWFGTGLVEELPACODIQLGLSLSLGLVGVPGVGLCYNALLVLANTLHASKASMT 60
DB 1 MMSCSWFGTGLVEELPACODIQLGLSLSLGLVGVPGVGLCYNALLVLANTLHASKASMT 60
QY 61 MPDYVFNMAVAGVLSALAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
DB 61 MPDYVFNMAVAGVLSALAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYST 120
QY 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSPSSILFYICSHVSTRALBCK 180
DB 121 ALLSDHYTERALPRTYMASVYNTRHVCGFWGALLTSPSSILFYICSHVSTRALBCK 180
QY 181 MONEAADAATLVFIVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHRLLVAATVC 240
DB 181 MONEAADAATLVFIVVPALATLVALLSRVRREDTPLDRDTGRLEPSAHRLLVAATVC 240
QY 241 TQFGMTPHYLLILGHTVYIISRGKPYDAHYIGLHFVDFSKLAFSSSFVTPPLLYRYMN 300
DB 241 TQFGMTPHYLLILGHTVYIISRGKPYDAHYIGLHFVDFSKLAFSSSFVTPPLLYRYMN 300
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 3
US-09-990-940-4
;; Sequence 4, Application US/09990940
;; Publication No. US20030027252A1
;; GENERAL INFORMATION:
;; APPLICANT: Tian, Hui
;; APPLICANT: Zhao, Jiayang
;; APPLICANT: Chen, Jin-Long
;; APPLICANT: Cutler, Gene
;; APPLICANT: An, Songzhu
;; APPLICANT: Dai, Kang
;; APPLICANT: Gupta, Jamila S.
;; APPLICANT: Tularek Inc.
;; TITLE OF INVENTION: No. US20030027252A1el Receptors
;; FILE REFERENCE: 018781-007410US
;; CURRENT APPLICATION NUMBER: US/09/990,940
;; CURRENT FILING DATE: 2001-11-21
;; PRIOR APPLICATION NUMBER: US 60/252,841
;; PRIOR FILING DATE: 2000-11-22
;; PRIOR APPLICATION NUMBER: US 60/257,636


```

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/261,377
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/279,554
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/280,696
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR339
US-09-990-940-4

Query Match      100.0%; Score 1733; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 9,8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MMSCSWNGTGLVEBELPACODLQIGLSLSLGLVGVPGVGLCYNALLVTIANHSKASMT 60
        1  MMSCSWNGTGLVEBELPACODLQIGLSLSLGLVGVPGVGLCYNALLVTIANHSKASMT 60
DB
QY      61  MPDYVFPVNMVAVGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSILVANYST 120
        61  MPDYVFPVNMVAVGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSILVANYST 120
DB
QY      121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
        121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
DB
QY      121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
        121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
DB
QY      181  MQNAEADATLVFIQVVPALATLVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
        181  MQNAEADATLVFIQVVPALATLVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
DB
QY      241  TQFGLMTPHYLLILGHTVIIISRGKPDVAHYLGHLFPKDSKLLAFSSFTVTPLLRYNM 300
        241  TQFGLMTPHYLLILGHTVIIISRGKPDVAHYLGHLFPKDSKLLAFSSFTVTPLLRYNM 300
DB
QY      301  QSFPSKLGRLMKKLPCCDRHCSPDHMGVQOVLA 333
        301  QSFPSKLGRLMKKLPCCDRHCSPDHMGVQOVLA 333
DB

RESULT 4
US-09-995-225-2
; Sequence 2, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huong T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Priddy, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9- Endogenous Versions of Human G
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
```

```

; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 2
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-2

Query Match      100.0%; Score 1733; DB 10; Length 333;
Best Local Similarity 100.0%; Pred. No. 9,8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MMSCSWNGTGLVEBELPACODLQIGLSLSLGLVGVPGVGLCYNALLVTIANHSKASMT 60
        1  MMSCSWNGTGLVEBELPACODLQIGLSLSLGLVGVPGVGLCYNALLVTIANHSKASMT 60
DB
QY      61  MPDYVFPVNMVAVGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSILVANYST 120
        61  MPDYVFPVNMVAVGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSILVANYST 120
DB
QY      61  MPDYVFPVNMVAVGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSILVANYST 120
        61  MPDYVFPVNMVAVGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSILVANYST 120
DB
QY      121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
        121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
DB
QY      121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
        121  ALLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTSFSSLLFYICSHVSTRALECAK 180
DB
QY      181  MQNAEADATLVFIQVVPALATLVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
        181  MQNAEADATLVFIQVVPALATLVLLSRVREDTPLDRDTGRLEPSAHRLLVATVC 240
DB
QY      241  TQFGLMTPHYLLILGHTVIIISRGKPDVAHYLGHLFPKDSKLLAFSSFTVTPLLRYNM 300
        241  TQFGLMTPHYLLILGHTVIIISRGKPDVAHYLGHLFPKDSKLLAFSSFTVTPLLRYNM 300
DB
QY      301  QSFPSKLGRLMKKLPCCDRHCSPDHMGVQOVLA 333
        301  QSFPSKLGRLMKKLPCCDRHCSPDHMGVQOVLA 333
DB

RESULT 5
US-10-165-844-13
; Sequence 13, Application US/10165844
; Publication No. US20030017539A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Hodges, Martin J.
; APPLICANT: Hunter, John J.
; APPLICANT: Rudolph-Owen, Laura
; APPLICANT: Welch, Nadine S.
; APPLICANT: Siles-Santego, Immaculada
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Encoding
; FILE REFERENCE: 35800/248302
; CURRENT APPLICATION NUMBER: US/10/165,844
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 09/088,857
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: US 09/324,465
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 09/464,685
; PRIOR FILING DATE: 1999-12-16
```

PRIOR APPLICATION NUMBER: US 09/741,783
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: US 09/145,745
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: US 09/383,745
PRIOR FILING DATE: 1999-08-26
PRIOR APPLICATION NUMBER: US 09/234,923
PRIOR FILING DATE: 1999-01-21
PRIOR APPLICATION NUMBER: US 09/340,880
PRIOR FILING DATE: 1999-06-28
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Faeseq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-165-844-13

Query Match 100.0%; Score 1733; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 9.8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVGVGLCYNALLVLANLHSAKSM 60
DB 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVGVGLCYNALLVLANLHSAKSM 60
QY 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALLSDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
QY 181 MONAERADATLVFVGVVPALATLVALLSRVRREDTPLDRDGRLEPSAHRLLVATVC 240
DB 181 MONAERADATLVFVGVVPALATLVALLSRVRREDTPLDRDGRLEPSAHRLLVATVC 240
QY 241 TOFGMLTPHYLLILGHTVYIISRGKPVDAHYLGHLHFVKDFSKLAFSSSFVTPLLYRYMN 300
DB 241 TOFGMLTPHYLLILGHTVYIISRGKPVDAHYLGHLHFVKDFSKLAFSSSFVTPLLYRYMN 300
QY 301 QSFPSKLGRLMKKLP CGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKLGRLMKKLP CGDRHCSPDHMGVOQVLA 333

RESULT 6
US-10-101-148-3
Sequence 3, Application US/10101148
Publication No. US20030187198A1
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Yeung, George
TITLE OF INVENTION: A No. US20030187198A1 Chemokine Receptor Obtained From a CDNA I
FILE REFERENCE: 28110/38317
CURRENT APPLICATION NUMBER: US/10/101,148
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-101-148-3

Query Match 100.0%; Score 1733; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 9.8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVGVGLCYNALLVLANLHSAKSM 60
DB 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVGVGLCYNALLVLANLHSAKSM 60

DB 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVGVGLCYNALLVLANLHSAKSM 60
QY 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALLSDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
QY 181 MONAERADATLVFVGVVPALATLVALLSRVRREDTPLDRDGRLEPSAHRLLVATVC 240
DB 181 MONAERADATLVFVGVVPALATLVALLSRVRREDTPLDRDGRLEPSAHRLLVATVC 240
QY 241 TOFGMLTPHYLLILGHTVYIISRGKPVDAHYLGHLHFVKDFSKLAFSSSFVTPLLYRYMN 300
DB 241 TOFGMLTPHYLLILGHTVYIISRGKPVDAHYLGHLHFVKDFSKLAFSSSFVTPLLYRYMN 300
QY 301 QSFPSKLGRLMKKLP CGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKLGRLMKKLP CGDRHCSPDHMGVOQVLA 333

RESULT 7
US-10-292-798-1016
Sequence 1016, Application US/10292798
Publication No. US20030235833A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1016
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-10-292-798-1016

Query Match 100.0%; Score 1733; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 9.8e-153;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVGVGLCYNALLVLANLHSAKSM 60
DB 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVGVGLCYNALLVLANLHSAKSM 60
QY 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVNMAVAGLVLSALAPVHLLGPPSSRWALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
DB 121 ALLSDHYIERALPRTYMASVYNTTRHVCGFVWGALLTSFSSLLFYICSHVSTRALECAK 180
QY 181 MONAERADATLVFVGVVPALATLVALLSRVRREDTPLDRDGRLEPSAHRLLVATVC 240
DB 181 MONAERADATLVFVGVVPALATLVALLSRVRREDTPLDRDGRLEPSAHRLLVATVC 240
QY 241 TOFGMLTPHYLLILGHTVYIISRGKPVDAHYLGHLHFVKDFSKLAFSSSFVTPLLYRYMN 300
DB 241 TOFGMLTPHYLLILGHTVYIISRGKPVDAHYLGHLHFVKDFSKLAFSSSFVTPLLYRYMN 300
QY 301 QSFPSKLGRLMKKLP CGDRHCSPDHMGVOQVLA 333
DB 301 QSFPSKLGRLMKKLP CGDRHCSPDHMGVOQVLA 333

Db 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333

RESULT 8

US-10-296-115-1174
; Sequence 1174, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1174
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1174

Query Match 100.0%; Score 1733; DB 15; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.1e-152;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVTANLHSAKSM 60
Db 38 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVTANLHSAKSM 97
QY 61 MPDYTFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
Db 98 MPDYTFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 157
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCQFWGALLTSFSSILFYICSHVSTRALECAK 180
Db 158 ALLSLDHYIERALPRTYMASVYNTTRHVCQFWGALLTSFSSILFYICSHVSTRALECAK 217
QY 181 MONAEADATLVITIGYVVPALATLVYALVLSRVRREDTPLDRDTGRLEPSAHLVAATVC 240
Db 218 MONAEADATLVITIGYVVPALATLVYALVLSRVRREDTPLDRDTGRLEPSAHLVAATVC 277
QY 241 TOFGLWTPHYLLIGHTVIISRGKPYDAHYLGILHFVKDFSKLLAFSSSVTPTLLYRYNM 300
Db 278 TOFGLWTPHYLLIGHTVIISRGKPYDAHYLGILHFVKDFSKLLAFSSSVTPTLLYRYNM 337
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
Db 338 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 370

RESULT 9
US-10-017-161-1202

; Sequence 1202, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1202
; LENGTH: 386

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-1202

Query Match 100.0%; Score 1733; DB 14; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.2e-152;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVTANLHSAKSM 60
Db 54 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVTANLHSAKSM 113
QY 61 MPDYTFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
Db 114 MPDYTFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 173
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCQFWGALLTSFSSILFYICSHVSTRALECAK 180
Db 174 ALLSLDHYIERALPRTYMASVYNTTRHVCQFWGALLTSFSSILFYICSHVSTRALECAK 233
QY 181 MONAEADATLVITIGYVVPALATLVYALVLSRVRREDTPLDRDTGRLEPSAHLVAATVC 240
Db 234 MONAEADATLVITIGYVVPALATLVYALVLSRVRREDTPLDRDTGRLEPSAHLVAATVC 293
QY 241 TOFGLWTPHYLLIGHTVIISRGKPYDAHYLGILHFVKDFSKLLAFSSSVTPTLLYRYNM 300
Db 294 TOFGLWTPHYLLIGHTVIISRGKPYDAHYLGILHFVKDFSKLLAFSSSVTPTLLYRYNM 353
QY 301 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 333
Db 354 QSFPSKIQRLMKKLPCGDRHCSPDHMGVOQVLA 386

RESULT 10

US-09-895-686-3
; Sequence 3, Application US/09895686
; Patent No. US20020106655A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN GPCR PROTEINS
; FILE REFERENCE: PC-0044 CIP
; CURRENT APPLICATION NUMBER: US/09/895,686
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020106655A1 2214673CD1
US-09-895-686-3

Query Match 99.6%; Score 1726; DB 9; Length 333;
Best Local Similarity 99.7%; Pred. No. 4.4e-152;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVTANLHSAKSM 60
Db 1 MMSCSWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALLVTANLHSAKSM 60
QY 61 MPDYTFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
Db 61 MPDYTFVNNAVAGLVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFVSSLVAMYST 120
QY 121 ALLSLDHYIERALPRTYMASVYNTTRHVCQFWGALLTSFSSILFYICSHVSTRALECAK 180
Db 121 ALLSLDHYIERALPRTYMASVYNTTRHVCQFWGALLTSFSSILFYICSHVSTRALECAK 180

QY 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHLVAVTC 240
DB 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHLVAVTC 240
QY 241 TOFGMTPHYLILGHTVVISRGPVDVAHYGLHFVWDFSKLAFSSFTPLLYRYMN 300
DB 241 TOFGMTPHYLILGHTVVISRGPVDVAHYGLHFVWDFSKLAFSSFTPLLYRYMN 300
QY 301 QSPFSKLORLMKKLPCCGRHCSPDHMGVQOYLA 333
DB 301 QSPFSKLORLMKKLPCCGRHCSPDHMGVQOYLA 333
RESULT 11
US-09-791-932-91
Sequence 91, Application US/09791932
Publication No. US20030003451A1
GENERAL INFORMATION:
APPLICANT: Vogeli, Gabriel
APPLICANT: Parodi, Luis A.
APPLICANT: Hiebsch, Ronald R.
APPLICANT: Lind, Peter
APPLICANT: Kaytes, Paul S.
APPLICANT: Ruff, Valerie
APPLICANT: Huf, Rita M.
APPLICANT: Wood, Linda S.
TITLE OF INVENTION: No. US20030003451A1 G Protein-Coupled Receptors Cross-Referen
FILE REFERENCE: 00325 US1
CURRENT APPLICATION NUMBER: US/09/791,932
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 60/184,305
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,304
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,303
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,397
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,247
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/188,880
PRIOR FILING DATE: 2000-03-13
PRIOR APPLICATION NUMBER: 60/217,369
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/217,370
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/218,492
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: 60/186,810
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/188,064
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 60/186,457
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: 60/213,861
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/194,344
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: 60/218,337
PRIOR FILING DATE: 2000-07-14
NUMBER OF SEQ ID NOS: 184
SOFTWARE: PatentIn version 3.0
SEQ ID NO 91
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-932-91
Query Match 99.6%; Score 1726; DB 10; Length 333;
Best Local Similarity 99.7%; Pred. No. 4,4e-152;
Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MMSCSWFGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNNALVLANIHSKASMT 60

DB 1 MMSCSWFGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNNALVLANIHSKASMT 60
QY 61 MPDYVFVMAVAVGLVLSALAPVHLGPPSSRWALMSVGGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVMAVAVGLVLSALAPVHLGPPSSRWALMSVGGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYIERALPRTYMASVYNTRHVCGFWMGALLTSFSSILFYICSHVSTRALBCKAK 180
DB 121 ALLSDHYIERALPRTYMASVYNTRHVCGFWMGALLTSFSSILFYICSHVSTRALBCKAK 180
QY 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHLVAVTC 240
DB 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHLVAVTC 240
QY 241 TOFGMTPHYLILGHTVVISRGPVDVAHYGLHFVWDFSKLAFSSFTPLLYRYMN 300
DB 241 TOFGMTPHYLILGHTVVISRGPVDVAHYGLHFVWDFSKLAFSSFTPLLYRYMN 300
QY 301 QSPFSKLORLMKKLPCCGRHCSPDHMGVQOYLA 333
DB 301 QSPFSKLORLMKKLPCCGRHCSPDHMGVQOYLA 333

RESULT 12
US-10-225-567A-680
Sequence 680, Application US/10225567A
Publication No. US20030113798A1
GENERAL INFORMATION:
APPLICANT: Lifespan Biosciences
APPLICANT: Brown, Joseph P.
APPLICANT: Burnet, Glenna C.
APPLICANT: Roush, Christine L.
TITLE OF INVENTION: ANTIBIOTIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: PatentIn version 3.1
SEQ ID NO 680
LENGTH: 333
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (11)..(11)
OTHER INFORMATION: Unknown Amino Acid
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (17)..(17)
OTHER INFORMATION: Unknown Amino Acid
US-10-225-567A-680

Query Match 99.1%; Score 1717; DB 14; Length 333;
Best Local Similarity 99.4%; Pred. No. 3e-151;
Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MMSCSWFGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNNALVLANIHSKASMT 60
DB 1 MMSCSWFGTGLVEELPACODLQGLSLSLGLVGVGVGLCYNNALVLANIHSKASMT 60
QY 61 MPDYVFVMAVAVGLVLSALAPVHLGPPSSRWALMSVGGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFVMAVAVGLVLSALAPVHLGPPSSRWALMSVGGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYIERALPRTYMASVYNTRHVCGFWMGALLTSFSSILFYICSHVSTRALBCKAK 180
DB 121 ALLSDHYIERALPRTYMASVYNTRHVCGFWMGALLTSFSSILFYICSHVSTRALBCKAK 180
QY 181 MONEAADATLVFIGVVPALATLYALVLSRVREDTPLDRDGRLEPSAHLVAVTC 240

Db 181 MONEAADATLVFGYVAPALATLVALVLSRVREDTPLDRDGRLEPSARLLVATVC 240
Qy 241 TORLWTPHYLLIGHVILSRGPVDAHYGLHPKXDSKLAFFSSFTPLLRYMN 300
Db 241 TQFLWTPHYLLIGHVILSRGPVDAHYGLHPKXDSKLAFFSSFTPLLRYMN 300
Qy 301 QSPSKLQRLMKKLPCCDRHCSPDHMGVOVLA 333
Db 301 QSPSKLQRLMKKLPCCDRHCSPDHMGVOVLA 333

RESULT 13
US-09-371-900-38
Sequence 38, Application US/09371900
Patent No. US20020137700A1
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/371,900
FILING DATE: 11-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-371-900-38

Query Match 15.6%; Score 271, DB 9; Length 375;
Best Local Similarity 27.5%; Pred. No. 1,4e-16;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

Qy 8 NGTGLVEELPACODLQGLSLSLGLVGVVGLCTNALVLNLHLSKASMTWPDYFV 67
Db 44 NGTG---ELSEHQYVIGL-FLSLGLYIFLPFGVGNILIVNIFRKRMTTFDLYFI 99
Qy 68 NMAVAGLVISALAPVHLLGPPSSRMALMSVGEVH-----VALQIPF-NVSSIVAMYSTA 121
Db 100 NLAVALDILVADSLIEVF-----NLHERYDIADVLCFMSLFLRVNMYSSV 145

Qy 122 L-----LSLDHYERIALPRTYMASVYNTRH-----VCGFVWGALLTSFSLFYICSHYST 173
Db 146 FELTMSFDPYIT--ALARMRCSLFRTKHARLSGGLIW--MASVATLVPTAVYLQH 200
Qy 174 RALECANAQNEAADATLVFGYVVP--ALATVYALVLSRVREDTPLDRDGRLEP-- 229
Db 201 TDEACFCFADVRVQWLEVTIGFTVPAIIGLCYSILVRLVDAH-----RHKG-LRPFR 254
Qy 230 -SARLLVAVTCQFGIMTPHYLLIGHVILSRGP-----VDHYGLHHPVKD 279
Db 255 QKALRMILAVLVFVCMLEPENVISVH--LQRTQGAAPCKQSPFRAHPL-TGHIV-- 309

Qy 280 FSKLAFSSFTVPLRYVMOSEPSKLORLMKK---LPCDRHCSPDHMGVOVLA 332
Db 310 -NLAAPSNCMLPLISFLGETFRDRLRYIKTNLPALDRFC--HAALKAVI 360

RESULT 14
US-09-924-417-59
Sequence 59, Application US/09924417
Patent No. US20020142441A1
GENERAL INFORMATION:
APPLICANT: FALB, DEAN
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR
DISEASE
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/924,417
FILING DATE: 07-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,286
FILING DATE: 04-MAR-1998
APPLICATION NUMBER: 08/870,434
FILING DATE: 06-JUN-1997
APPLICATION NUMBER: 08/799,910
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: 60/011,787
FILING DATE: 16-FEB-1996
APPLICATION NUMBER: 08/599,654
FILING DATE: 09-FEB-1996
APPLICATION NUMBER: 08/485,573
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-114-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 7909090
TELEFAX: (212) 8699741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-924-417-59

Query Match 15.6%; Score 271; DB 9; Length 375;
Best Local Similarity 27.5%; Pred. No. 1.4e-16;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACQDLDLGSLSLGLVGVPGVCYNALLVLANLHSAKSMTPDVFV 67
DB 44 NGTG--ELSEHQQYVIGL-FLSCLYTIFLPIGVGNILILVNVISFREKKTIPDLFYI 99
QY 68 NNAVAGLVLSALAPVHLGPPSSRWALMSVSGEVH-----VALQIPF-NVSSLVAMYSTA 121
DB 100 NLAVALDLIVADSLIEVF-----NLHERYYDIAVLCFMSLFLRVNMYSSV 145
QY 122 L-----LSLDHYTERALPRTYMASVYNTRH---VCGFWGALLTSFSSLLFYICSHYST 173
DB 146 FLITWMSFDRYI--ALARMRCSLFRTGHARLSGLIM--MASVSATLVPTAVHLQH 200
QY 174 RALECAKQONAEADATIVFIGYVP--ALATYVALVLSVRREDTPLDRTGRLEP-- 229
DB 201 TDEACFCPADVREVOVMEVTLGFIIVPAIIGLCYSLIVRVRAH-----RHRG-LRP RR 254
QY 230 -SAHRLVATVCTOGMTPTPHYLILGHTVLIISRGK-----VDANYLGHLHFVKD 279
DB 255 QKALRMILAVALVVFVVCWLPENVFISVH--LLQRTQGAAPCKQSFRAHPL-TGHIV-- 309
QY 280 FSKLIAFSSFVTPLLYRYNMQSFPSKIQRLMKK---LPGDRHCSPDHNGVOQVL 332
DB 310 --NLAAFSNSCLNPLIYSLGFTPRDLRLYIEQKTNLPALDRFC--HAALKAVI 360

RESULT 15

US-09-970-820-38
Sequence 38, Application US/09970820
Patent No. US20020170077A1
GENERAL INFORMATION:

APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/970,820
FILING DATE: 05-Oct-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:

NAME: Cornuzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-032

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-970-820-38

Query Match 15.6%; Score 271; DB 9; Length 375;
Best Local Similarity 27.5%; Pred. No. 1.4e-16;
Matches 98; Conservative 64; Mismatches 124; Indels 70; Gaps 19;

QY 8 NGTGLVEELPACQDLDLGSLSLGLVGVPGVCYNALLVLANLHSAKSMTPDVFV 67
DB 44 NGTG--ELSEHQQYVIGL-FLSCLYTIFLPIGVGNILILVNVISFREKKTIPDLFYI 99
QY 68 NNAVAGLVLSALAPVHLGPPSSRWALMSVSGEVH-----VALQIPF-NVSSLVAMYSTA 121
DB 100 NLAVALDLIVADSLIEVF-----NLHERYYDIAVLCFMSLFLRVNMYSSV 145
QY 122 L-----LSLDHYTERALPRTYMASVYNTRH---VCGFWGALLTSFSSLLFYICSHYST 173
DB 146 FLITWMSFDRYI--ALARMRCSLFRTGHARLSGLIM--MASVSATLVPTAVHLQH 200
QY 174 RALECAKQONAEADATIVFIGYVP--ALATYVALVLSVRREDTPLDRTGRLEP-- 229
DB 201 TDEACFCPADVREVOVMEVTLGFIIVPAIIGLCYSLIVRVRAH-----RHRG-LRP RR 254
QY 230 -SAHRLVATVCTOGMTPTPHYLILGHTVLIISRGK-----VDANYLGHLHFVKD 279
DB 255 QKALRMILAVALVVFVVCWLPENVFISVH--LLQRTQGAAPCKQSFRAHPL-TGHIV-- 309
QY 280 FSKLIAFSSFVTPLLYRYNMQSFPSKIQRLMKK---LPGDRHCSPDHNGVOQVL 332
DB 310 --NLAAFSNSCLNPLIYSLGFTPRDLRLYIEQKTNLPALDRFC--HAALKAVI 360

Search completed: December 14, 2004, 20:27:30
Job time : 47.606 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 20:04:39 ; Search time 14.5627 Seconds
(without alignments)
2200.157 Million cell updates/sec

Title: US-09-995-225B-2

Perfect score: 1733

Sequence: 1 MWCSEWNGNGVLEBPAQ.....LPGGDHCHSPDHNGVQVLA 333

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR 79: *
2: PIR: *
3: PIR: *
4: PIR: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	267	15.4	375	2 JCS069	G protein-coupled
2	240.5	13.9	358	2 G02670	IL8-related recept
3	233.5	13.5	375	2 JCS509	G protein-coupled
4	205	11.8	358	2 A53752	interleukin-8 rece
5	200	11.5	350	2 A53611	interleukin-8 rece
6	193.5	11.2	350	2 A53945	interleukin-8 rece
7	185	10.7	355	2 JQ1231	interleukin-8 rece
8	184.5	10.6	369	2 JCS068	G protein-coupled
9	178.5	10.3	378	2 A55735	G protein-coupled
10	177.5	10.2	372	2 B55735	lymphocyte-specifi
11	177	10.2	372	2 S26657	G protein-coupled
12	175.5	10.1	367	2 JCS049	interferon-inducib
13	169	9.8	399	2 S29480	homobesin recepto
14	167	9.6	327	2 S56162	MCCR15 protein - h
15	167	9.6	354	2 I58186	probable G protein
16	166.5	9.6	352	2 I38973	G protein-coupled
17	166.5	9.6	352	2 A45747	neuropeptide Y/pep
18	166	9.6	359	2 A48921	interleukin-8 rece
19	165.5	9.5	378	2 A45680	somatostatin recep
20	165.5	9.5	388	2 JN0605	somatostatin recep
21	165.5	9.5	423	2 JCT677	allatostatin recep
22	164.5	9.5	353	2 S28787	neuropeptide Y/pep
23	164.5	9.5	359	2 A42656	angiotensin II rec
24	163	9.4	352	2 JCS0296	thyrotropin releas
25	162.5	9.4	384	2 A47249	brain-specific som
26	162	9.3	374	2 S42628	G protein-coupled
27	161.5	9.3	352	2 G00048	fusin (LESTRA) - c
28	161.5	9.3	354	2 A23659	interleukin-8 rece
29	161.5	9.3	356	2 S42096	interleukin-8 rece

30	161.5	9.3	359	2 JQ1516	angiotensin II rec
31	161.5	9.3	418	2 A46226	somatostatin recep
32	161	9.3	374	2 S32785	G protein-coupled
33	161	9.3	399	2 A46632	homobesin-like pept
34	160.5	9.3	366	2 S71152	neuropeptide Y/pep
35	160	9.2	355	2 A55733	G protein-coupled
36	159.5	9.2	391	2 A39297	somatostatin recep
37	158.5	9.1	363	2 I57940	somatostatin recep
38	158.5	9.1	391	2 C41795	somatostatin recep
39	158.5	9.1	428	2 S30508	probable G protein
40	157.5	9.1	428	2 A44021	somatostatin recep
41	155.5	9.0	384	2 JCS4629	somatostatin recep
42	155.5	9.0	391	2 A41795	somatostatin recep
43	155	8.9	362	2 I38990	Mel1b-melanconin re
44	155	8.9	362	2 A30341	G protein-coupled
45	154.5	8.9	350	2 JN0621	G protein-coupled

ALIGNMENTS

RESULT 1

JCS069

G protein-coupled receptor CMRL2 - human

N:Alternate names: constitutively expressed peptide-like receptor; flow-induced endothe

C:Species: Homo sapiens (man)

C>Date: 31-Jan-1997 #sequence, revision 28-Aug-1998 #ext_change 09-Jul-2004

C/Accession: JCS069; JCS786; JCS294

R/Owman, C.; Blay, P.; Nilsson, C.; Lolait, S.J.

A:Title: Cloning of human cDNA encoding a novel heptahelix receptor expressed in Burkitt

A:Reference number: JCS069; MID:97079175; PMID:8920907

A/Accession: JCS069

A/Molecule type: mRNA

A/Residues: 1-375 <OM>

A/Cross-references: UNIPROT:Q99527; GB:Y08162; MID:91707499; PIDN:CA69354.1; PID:91707

A/Experimental source: B-cell lymphoblast

R/Tkaka, Y.; Kato, C.; Kondo, S.; Korenaga, R.; Ando, J.

A:Title: Cloning of cDNAs encoding G protein-coupled receptor expressed in human endothe

A/Reference number: JCS786; MID:98063308; PMID:9398636

A/Accession: JCS786

A/Molecule type: mRNA

A/Residues: 1-375 <TK>

A/Cross-references: GB:AF015257; MID:92353152; PIDN:AAC51904.1; PID:92353153

A/Experimental source: umbilical vein endothelial cell

R/Feng, Y.; Gregor, P.

A:Title: Cloning of a novel member of the G protein-coupled receptor family related to

A/Reference number: JCS294; MID:97224403; PMID:9070864

A/Accession: JCS294

A/Molecule type: DNA

A/Residues: 1-311, 'T', 313-375 <FEN>

A/Cross-references: GB:U77827; MID:91906591; PIDN:AAC51173.1; PID:91906592

C/Comment: This protein plays a role in B-cell functions and is involved in endothelial

C/Genetics:

A/Gene: GDB:CMKRL2; CEPR

A/Cross-references: GDB:3929190

C/Superfamily: vertebrate rhodopsin

F/64-83/Domain: G protein-coupled receptor; glycoprotein; transmembrane protein

F/96-114/Domain: transmembrane #status predicted <TM1>

F/126-148/Domain: transmembrane #status predicted <TM2>

F/160-198/Domain: transmembrane #status predicted <TM3>

F/221-238/Domain: transmembrane #status predicted <TM4>

F/260-283/Domain: transmembrane #status predicted <TM5>

F/308-327/Domain: transmembrane #status predicted <TM6>

F/32,32,44/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.4%; Score 267; DB 2; Length 375;

Best Local Similarity 27.4%; Pred. No. 1,3e-14;

Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;

```

QY      8 NGTGVLEELPACDQIQQLSLSLSLGLVAVGVGLCTYAAVLVLANHASKAMTPDYFV  67
Db      44 NGTG---ELSEHQYVIGL-FLSCLYTIFLPPIGVGNILLVNNISFEREKMTIPDLYFI  99
QY      68 NMAVAGVLSALAPRHHLLGPPSSRWALMSVGGEVHVALQIPNNISLVAMRSTL-----L  123
Db      100 NMAVADLLVADSLLEFVNLHERYYDIAVLCTFMSLFQ-----VNMSSVFFLTMM  151
QY      124 SLDHYTEALPRTYMAVYNTTRH-----VCGFVWGALLTSPSSLFYICSHVSTRALECA  179
Db      152 SPDRKI-ALBARARCSLFRKTHARLSGGLIW--MASVATLVPPFAVHVLQHTDEACF  206
QY      180 KMQNAEADATLVFIFYVVP--ALATLVAVLVLSVRREDPTLDRDGTGLEP--SAHRL  234
Db      207 CFADVREVQWLEVTLGFIVPFALIGCYSLVRLVLRNH-----RHNG-LRPFRQKALRM  260
QY      235 LVATYCTQFGLMTPHYLILLGHTVYIISRGP-----VDAYHGLGHFVKDFSKLLA  285
Db      261 ILAVLVFVFCWMLPENVFISVH--LQRTQGAAPCKQSPFHAHPL-TGHTIV---NLAA  313
QY      286 FSSSFVTEFLKRYMNSQSPSKQLQMLMK---LPGCDRHCSPDHNGVQVOYL  332
Db      314 FSNSGCLNLIYSFGETFRDKLRLYIEQKTNLPLNRC--NAALRAVI  360

```

```

RESULT 2
G02670
IL8-related receptor - human
C|Species: Homo sapiens (man)
C|Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C|Accession: G02670
R|Mccoy, R.L., Perlmuter, D.H.
submitted to the EMBL Data Library, May 1996
A|Reference number: H01575
A|Accession: G02670
A|Status: preliminary; translated from GB/EMBL/DDJ
A|Molecule type: mRNA
A|Residues: 1358 <MCC>
A|Cross-references: UNIPROT:Q99527, EMBL:U58828, NID:g1381668, PIDN:AAB02736.1, PID:g138
C|Genetics:
A|Gene: DRVY12
C|Superfamily: vertebrate rhodopsin

Query March          13.9%; Score 240.5; DB 2; Length 358;
Best Local Similarity 27.2%; Pred. No. 2e-12;
Matches 90; Conservative 61; Mismatches 131; Indels 49; Gaps 16;

Qy      15 ELPAQODIQQLSLSTLSLTGLVVGVPVGCYALLVLNLNTSKASMTMPDYVFVNMAVAGL 74
       |||||:::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      30 ELSHQQGVIGL-FLSCLYTIFLPFGVGNILLIVNISREKMTTPDLXFTINLAVDL 88

Qy      75 VLSALAPVHLIGPPSSRMALMSVGGEVHALQIPFNVSILVAMYSTAL--LSLDHYTERA 132
       |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      89 ILVADSLIEVFNLERYYDIALVLCFMSLFLOQH-----VQAASSFLTWMSFDRIY--A 141

Qy      133 LPRITMASVNNTRH----VCGFWNGCALITSFSLLP-YICSHYSTALCAEQNAEEA 167
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      142 LARMRCSLFEPTKHARLSCLTIW---MSSVSATLVPFRAVHLQHTDEACFCPADREV 197

Qy      188 DATLVFIGVYP--ALATLVALLVLSRVREDTFLDBDTGRLEP---SAHRILVAATCTQ 242
       |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      198 QWLSEVTIIGFIYPAIIIGCYSLIRVLVRAN-----RHKG-LRRRKAKALRMILAIVLV 251

Qy      243 FGLMTPEYLIILGHTVIISRGKP-----VDARYLGILHFVKDFSKLIASSSFVTP 293
       |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      252 FVCMLPEWVPSIVH--LIQRTQPGAAPCKOSFRHAHPL-TGHIV---NLAAFSNSCLTP 304

Qy      294 LLRYRNMOFPSKLQRLMKK---LPCGDRIHC 321
       |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      305 LIYSFLGTFPRDKRLTYIEQKTNPALNRFC 335

```

JC5509
 G protein-coupled receptor 41 - rat
 C.Species: Rattus norvegicus (Norway rat)
 C.Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 09-Jul-2004
 C.Accession: JC5509
 R.Bonini, J.A.; Anderson, S.M.; Steiner, D.F.,
 Biochem. Biophys. Res. Commun., 234, 190-193, 1997
 A.Title: Molecular cloning and tissue expression of a novel orphan G protein-coupled rec
 A.Reference number: JC5509; MUID:97312546; PMID:9168967
 A.Accession: JC5509
 A.Molecule type: mRNA
 A.Residues: 1-375 <BON>
 A.Cross-references: UNIPROT:O08878; GB:U92802; NID:g2138341; PIDN:AAC53208.1; PID:g21383
 A.Experimental source: lung
 C.Comment: This protein transduces the signal of a wide variety of hormones, neurotransm
 target cells.
 C.Superfamily: vertebrate rhodopsin
 C.Keywords: glycoprotein; phosphoprotein
 F.32/44,84/Binding site: carboxylate (asn) (covalent) #status predicted
 F.86/297/Binding site: phosphate (Ser) (covalent) #status predicted
 F.355/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match	13.5%	Score 233.5	DB 2	Length 375
Best Local Similarity	24.6%	Fred. No. 8,1e-12		
Matches	85	Conservative 69	Mismatches 138	Indels 53
				Gaps 16
QY	12	LVEELPACODIQLGLSLTSLGLVGVGVGVCYNALLVLANLHSAASMTMPDVFVFNMAV	71	
DB	45	LTGDISEQQVIAL-FLSCLYTIPLFPIGVGNILLIIVNISFPEKKTIIPDPIFINLAA	103	
QY	72	AGLVLSALAPVHLGPPSSRWALMSVGVEVHALQIPNVSLVAMYSTAL---LSLDH	127	
DB	104	ADLILVADSLIEVFNLDEQYDIADVLCFMSLFDQI-----NMYSVSFPLTMMSPDR	155	
QY	128	YIERALPPTWASVYNTNH---VCGFWGAGLLTSFSLFYICSHYSTALCEAKMION	183	
DB	156	YL-ALAAKMRCGEFTRGHARLSGGLIW---MASVASLTPVFTVHLRHTEACFCFAD	210	
QY	184	AEADATLVFIGYVP-ALATLVALVLSVRREDPLDSDTGLEP---SAHRLLVAT	238	
DB	211	VREVOMLEVTIGFIVPFAIIGLCTSLIRALIRAH---RHRC-LRPRRQALAMITPAV	264	
QY	239	VCTQFGMLTPHYLLLGHTVVISRGPVDA-----HYGLL-HFVXDPSKLLFSSGSF	290	
DB	265	VLVFFICLPENVFISVH-LLQMAQPEDPTCKQSFRAIAYPLTGHIIV---NLAAFPSNC	318	
QY	291	VTPLIYRYMNSFPSKIQRLMKK---LPCGRHCSGPDMAQVOVL	332	
DB	319	LSPIIYSELTGFTRDKRLVVAQKTSPLPALNRFC---HATLKAVI	360	

```

RESULT 4
A53752
Interleukin-8 receptor (clone 5B1a) - rabbit
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: A53752
R/Prdo: G.N.; Thomas, K.M.; Suzuki, H.; Lahosa, G.J.; Wilkinson, N.; Folco, E.; Navarrete
J. Biol. Chem. 269, 12391-12399, 1994
A/Title: Molecular characterization of a novel rabbit interleukin-8 receptor isoform.
A/Reference number: A53752; MUID:94230294; PMID:8175642
A/Accession: A53752
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-358 <PPA>
A/Cross-references: UNIPROT:P35344; GB:L24445; NID:G437661; PIDW:AAA33378.1; PID:G437662
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match      11.8%; Score 205; DB 2; Length 358;
Best Local Similarity 23.0%; Pred. No. 1.8e-09;
Matches 70; Conservative 58; Mismatches 122; Indels 54; Gaps 9;

```


QY 28 LLSLLGLVGVPGVGLCYNALVLANTHSKASMTMPDVFYVNMVAGLVLSALAPVHLGP 87
 DB 58 LLSLLG-----NSLVMVLIIYSRSTCSTVDYLNALADLLFATLPI----- 101
 QY 88 PSSRMAL-----NSVGEVHALQIPFNVSLSVAMYSTALLSDHYIERA-LPRTYMAV 141
 DB 102 -----MAASKVHGMPFGPFLCKVSLKEVNFYSGLILLACISVDRIYLAIVHARTMIOKR 157
 QY 142 YNTRHVCGFVWGALITFSFSLFYICSHVSTRALBACAKOMAEADATLVF-----IG 195
 DB 158 HLKVFICLSIMMGVSLISLPIILFRNALFPFNSSPVCEDMGNSTAKRMVLRILPOTFG 217
 QY 196 YVVPALATLVALVLLSRVRBDPLDRDGLRLEPSAHRLLVAATVCTQFGMLTPHYLL 255
 DB 218 FILPLVLMFCYGFPTLR-----TLFQAHMQQ-KRRAMRVIPAVVLIPLCLMPLPYNLVLT 271
 QY 256 HTVI-----ISRGKPVDAHYIGLHPVDSKSLAFSSFPVTLVRYMNSPFSKL 307
 DB 272 DTLMTHTVIOETCERRANDIR-----ALDANTEILGFLHSCINPLIIVAFIQKFRYGL 323
 QY 308 QRLM 311
 DB 324 LKIL 327

RESULT 5

interleukin-8 receptor type B - human
 A:Species: Homo sapiens (man)
 C:Date: 07-Oct-1994 #sequence revision 12-Apr-1996 #text_change 09-Jul-2004
 C:Accession: I37898; I38712; A53611; A39446
 R:Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
 J. Biol. Chem. 269, 26381-26389, 1994
 A:Title: Comparison of the genomic organization and promoter function for human interleukin-8 receptor.
 A:Reference number: I37898; MUID:95014476; PMID:7929358
 A:Accession: I37898
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-360 <RES>
 A:Cross-references: UNIPROT:P25025; EMBL:U11869; NID:G511801; PIDN:AA60656.1; PID:G5118
 A:Accession: I38712
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-15 <RES>
 A:Cross-references: EMBL:U11872; NID:G511808; PIDN:AA64380.1; PID:G511809; EMBL:U11873; I1876; NID:G511816; PID:G511817; EMBL:U11877; NID:G511818; PID:G511819; EMBL:U11878; NID:J. Biol. Chem. 269, 11065-11072, 1994
 A:Title: Structure, genomic organization, and expression of the human interleukin-8 receptor.
 A:Reference number: A53611; MUID:94209273; PMID:7512557
 A:Accession: A53611
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 6-360 <SPR>
 A:Cross-references: GB:M99412; GB:L19593
 R:Murphy, P.M.; Tiffany, H.L.
 Science 253, 1280-1283, 1991
 A:Title: Cloning of complementary DNA encoding a functional human interleukin-8 receptor.
 A:Reference number: A39446; MUID:91368200; PMID:1891716
 A:Accession: A39446
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 6-360 <MOR>
 A:Cross-references: GB:M73969
 C:Comment: This receptor, unlike IL8RA, binds several peptides besides interleukin-8, in C:Genetic:
 A:Gene: GDB:IL8RA; IL8RA
 A:Cross-references: GDB:I27868; OMIM:146928
 A:Map position: 2q35-2q35
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein
 Query Match 11.5%; Score 200; DB 2; Length 360;

Best Local Similarity 24.4%; Pred. No. 4,6e-09;
 Matches 76; Conservative 54; Mismatches 113; Indels 68; Gaps 11;

QY 28 LLSLLGLVGVPGVGLCYNALVLANTHSKASMTMPDVFYVNMVAGLVLSALAPVHLGP 87
 DB 60 LLSLLG-----NSLVMVLIIYSRGRSVTDVYLNALADLLFATLPI----- 103
 QY 88 PSSRMALMSVGEVH-----VALQIPFNVSLSVAMYSTALLSDHYIERA-LPRTYMA 139
 DB 104 -----MAASKVNGMIFGTPFLCKVSLKEVNFYSGLIL--LACISVDRIYLAIVHARTLITQ 157
 QY 140 SVNTRHVCGFVWGALITFSFSLFYICSHVSTRALBACAKOMAEADATLVF----- 193
 DB 158 KRLVVKFICLSITGSLILALPVLFRRTYSSNVSPACEDMGNTAMRMILRLIPQS 217
 QY 194 IGVVPALATLVALVLLSRVRBDPLDRDGLRLEPSAHRLLVAATVCTQFGMLTPHYLL 253
 DB 218 FGIVPLIIMLFCYGFPTLR-----TLFKAHMQQ-KRRAMRVIPAVVLIPLCLMPLPYNLVLT 271
 QY 254 LGHTVIISR-----GKPVDA-HYGLHFKVDSKSLAFSSFPVTLVRYMN 300
 DB 272 LADTLMTQVIOETCERRRHIDRALDATTILGILH-----SCLNPLIYAFIG 318
 QY 301 QSFPSKQRLM 311
 DB 319 QKFRHGLKIL 329

RESULT 6

A39445
 interleukin-8 receptor type A - human
 N:Alternate names: interleukin-8 receptor, high-affinity
 C:Species: Homo sapiens (man)
 C:Date: 22-Jan-1993 #sequence revision 12-Apr-1996 #text_change 09-Jul-2004
 C:Accession: I37449; I38710; I38711; A39445
 R:Mollereau, C.; Muscatelli, F.; Mattei, M.G.; Vassart, G.; Parmentier, M.
 Genomics 16, 248-251, 1993
 A:Title: The high-affinity interleukin-8 receptor gene (IL8RA) maps to the 2q33-q36 reg
 A:Reference number: I37449; MUID:93252387; PMID:8486366
 A:Accession: I37449
 A:Molecule type: DNA
 A:Residues: 1-350 <RES>
 A:Cross-references: UNIPROT:P25024; EMBL:X65858; NID:G312046; PIDN:CAA46688.1; PID:G312
 A:Residues: 1-350 <RES>
 A:Cross-references: EMBL:U11870; NID:G511804; PIDN:AA64378.1; PID:G511805
 A:Accession: I38711
 A:Molecule type: mRNA
 A:Residues: 1-16 <RES>
 A:Cross-references: EMBL:U11871; NID:G511806; PIDN:AA64379.1; PID:G733002
 R:Holmes, W.E.; Lee, J.; Kang, W.J.; Rice, G.C.; Wood, W.I.
 Science 253, 1278-1280, 1991
 A:Title: Structure and functional expression of a human interleukin-8 receptor.
 A:Reference number: A39445; MUID:91368199; PMID:1840701
 A:Accession: A39445
 A:Molecule type: mRNA
 A:Residues: 1-275, "T", 277-350 <HOL>
 A:Cross-references: GB:M68932; NID:G186369; PIDN:AA59159.1; PID:G186370
 C:Genetic:
 A:Gene: GDB:IL8RA
 A:Cross-references: GDB:I35039; OMIM:146929
 A:Map position: 2q35-2q35
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein
 Query Match 11.2%; Score 193.5; DB 2; Length 350;
 Best Local Similarity 23.9%; Pred. No. 1.5e-08;
 Matches 84; Conservative 54; Mismatches 117; Indels 97; Gaps 14;

```

QY      1  MMSGSRNGGGLVEELPACDLO-----LGLSLSLGLGVGVGL 42
      |||
      9  MMDPDLNFTGMP---PADEYSPQMLETETANKRYVILVAVLAFSLILG----- 56
      |||
QY      43  CYNALVLVLANHSAKMTMPDVFVNNVAGLVLSALAPVHLGPPSSRMALMSVGEVH 102
      |||
      57  --NSLVNLVILYSRVSVDVYLNLALADLFLALFLPT-----MAASVNGWIF 105
      |||
QY      103 VALQIPFNVSLSL---VAMYSTALLSLDHYIERALPRTYMASVYNT-----RHVCGFV- 151
      |||
      106 GTFPLC--KVVSLKEVNFYSGIILLACTISYDR-----YLAIVHARTRLTGKRLHVKFVCL 158
      |||
QY      152 --MGALLTSSSLLEFICGSHVSTRALCAKQMAEAADATVP-----IGVYVPLAT 203
      |||
      159 GCMGLSNVLSLPEFLFRQAHYNNSSPVCEVGLGNDPAKRMVRLILPHTFGFTVPLEFM 218
      |||
QY      204 LVALLVLSRVREDTPDLDRDTGRLEPSAHLVAVTCTQFGMLTWPVHLILGHVITSR- 262
      |||
      219 LFCYGFLLR-----TLFKAHMGQ-KHRAKRVIPAVVLIIFLLCMLPYNVLADTLMTQY 272
      |||
QY      263 -----GKRVDAHYGLHFVKDFSKLAFSSSFYTPLLRYMNSF 303
      |||
      273 IOBSCERRNNIGRALDA-----TEILGFLHSCINPIIYAFIQNF 312
      |||
RESULT 7
J01231
Interleukin-8 receptor - rabbit
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C/Accession: J01231; A46483
R/Beckmann, M.P.; Munger, W.E.; Kozlosky, C.; Vandenbos, T.; Price, V.; Lyman, S.; Gerard
Biochem. Biophys. Res. Commun. 179, 784-789, 1991
A/Title: Molecular characterization of the interleukin-8 receptor.
A/Reference number: J01231; MUID:91378994; PMID:1898400
A/Accession: J01231
A/Molecule type: DNA
A/Residues: 1-355 <BEC>
A/Cross-references: UNIPROT:P21109; GB:M74240; NID:g165438; PIDN:AAA31375.1; PID:g165439
R/Lee, J.; Kang, W.J.; Rice, G.C.; Wood, W.I.
J. Immunol. 148, 1261-1264, 1992
A/Title: Characterization of complementary DNA clones encoding the rabbit IL-8 receptor.
A/Reference number: A46483; MUID:92148149; PMID:1737938
A/Accession: A46483
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-355 <LEE>
A/Cross-references: GB:M82873; NID:g165440; PIDN:AAA31376.1; PID:g165441
A/Experimental source: neutrophils
A/Note: sequence extracted from NCBI backbone (NCBI:81526, NCBI:81530)
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match      10.7%; Score 185; DB 2; Length 355;
Best Local Similarity 21.9%; Pred. No. 8e-08;
Matches 77; Conservative 65; Mismatches 133; Indels 76; Gaps 14;

QY      1  MMSGSRNF-----NGTGL--VEE-----LPACDLO-----LGLSLSLGLGVGV 40
      |||
      11  LM--TWFEDEPAATATMPPEKDYSPCLVVTQTLNKYVVVYVAVLAFSLILG----- 61
      |||
QY      41  GLCYNALVLVLANHSAKMTMPDVFVNNVAGLVLSALAPVHLGPPSSRMAL-----W 95
      |||
      62  ----NSLVNLVILYSRVSVDVYLNLALADLFLALFLPT-----MAVSREKGM 108
      |||
      96  SVGGEVVALQIPFNVSLSLVAMYSTALLSLDHYIERA-LPRTYMASVYNTTRHVGCFVWG 154
      |||
      109 IGTPLCKVVSIVKEVNFYSGIILLACTISYDRILVHARTRLTGKRLHVKFICIGIAML 168
      |||
QY      155 ALLTSSSLLEFICGSHVSTRALCAKQMAEAADATVP-----IGVYVPLATLYALV 208
      |||
      169 SLTSLPFLFRQVSPNNSSPVCEYEDLGHNTAKRMVRLILPHTFGFTLPLVNLFCYG 228
      |||

```

```

QY      209 LLSRVREDTPDLDRDTGRLEPSAHLVAVTCTQFGMLTWPVHLILGHVIT- 260
      |||
      229 FTLR-----TLFQAHMGQ-KHRAKRVIPAVVLIIFLLCMLPYNVLADTLMTHTVIQERC 282
      |||
QY      261 SRKRPDAHYGLHFVKDFSKLAFSSSFYTPLLRYMNSFPSPKQLM 311
      |||
      283 QRNDRIDR-----ALDATILGFLHSCINPIIYAFIQNFRNGFLKML 325
      |||
RESULT 8
JCS068
G protein-coupled receptor CCR-13 - human
C/Species: Homo sapiens (man)
C/Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 21-Jul-2000
C/Accession: JCS068
R/Zaballos, A.; Varona, R.; Gutierrez, J.; Lind, P.; Marquez, G.
Biochem. Biophys. Res. Commun. 227, 846-853, 1996
A/Title: Molecular cloning and RNA expression of two new human chemokine receptor-like 9
A/Reference number: JCS067; MUID:97040707; PMID:8886020
A/Accession: JCS068
A/Molecule type: DNA
A/Residues: 1-369 <ZAB>
A/Cross-references: EMBL:Z79784; NID:g1668737; PIDN:CAB02144.1; PID:g1668738
C/Comment: This protein belongs to the family of alpha chemokine receptors.
C/Genetics:
A/Gene: GDB:CMKBR6; STRL22; GPR29; CCR6; CCR-13; GPR-CY4
A/Cross-references: GDB:5370639; OMIM:601835
A/Map position: 6q27-6q27
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein
F/42-68/Domain: transmembrane #status predicted <TM1>
F/79-99/Domain: transmembrane #status predicted <TM2>
F/115-136/Domain: transmembrane #status predicted <TM3>
F/160-180/Domain: transmembrane #status predicted <TM4>
F/212-233/Domain: transmembrane #status predicted <TM5>
F/250-271/Domain: transmembrane #status predicted <TM6>
F/292-315/Domain: transmembrane #status predicted <TM7>

Query Match      10.6%; Score 184.5; DB 2; Length 369;
Best Local Similarity 21.1%; Pred. No. 9.1e-08;
Matches 68; Conservative 59; Mismatches 104; Indels 91; Gaps 11;

QY      41  GLCYNALVLVLANHSAKMTMPDVFVNNVAGLVLSALAPVHLGPPSSRMALMSVGE 100
      |||
      55  GLGNILVITVFAFYKARSMTDVYLLMALADILFVLTLPFVAVSHATGAMVSNAICK 114
      |||
QY      101  VHAL-QIPFNVSLSLVAMYSTALLSLDHYI-----ERALPRTYMASVYNTTRHV 147
      |||
      115  LKGIYALNFCNML-----LTCISMRYIAIVQATSFRLRSKTLPR-----SKIT 162
      |||
QY      148  GCFVWGALLTSSSLLEFICGSHVSTRALCAKQMAEAADA----- 189
      |||
      163  CLVWGLSVIISSTFV-----NQKNYQGSVPCPKYQVSEPIRWKLM 209
      |||
QY      190  ---TLVFGYVVPALATL-YALVLSRVREDTPDLDRDTGRLEPSAHLVAVTCTQFG 244
      |||
      210  LGLELTF-GFTPLPMFMFTCYTFTVKTLYQAKNSRKH-----AIRVIAVAVLVELA 260
      |||
QY      245  LMTPEYLLILGHVILYSRGPVDAHYLG-----LHFVDFSKLAFSSSFYTP 294
      |||
      261  CQIHNNVNL-----VTANLGRKMSQSEKLTGYTIVTEVLAFLHCLINPV 309
      |||
QY      295  LYRINQSPFSKQLQRLMKQLPC 316
      |||
      310  LYAFIQKFRNYPFLKILKDLWC 331
      |||
RESULT 9
A53735
G protein-coupled receptor EB11 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004

```

C:Accession: A55735
R:Schweickart, V.L.; Raport, C.J.; Godiske, R.; Byers, M.G.; Eddy Jr., R.L.; Shows, T.B.
Genomic 23, 643-650, 1994
A:Title: Cloning of human and mouse EB1, a lymphoid-specific G-protein-coupled receptor
A:Reference number: A55735; MUID:95154835; PMID:7851893
A:Accession: A55735
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-378 <SC>
A:Cross-references: UNIPROT:P47774; GB:LJ1580; NID:g468340; PIDN:AAA74232.1; PID:g468344
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor

```

Query Match 3.0, 3.4; Score 178.5; DB 2; Length 378;
Best Local Similarity 23.9%; Pred. No. 3e-07;
Matches 73; Conservative 48; Mismatches 126; Indels 59; Gaps 10;

QY 40 VGLCYNALVLTIANHSKASMTMPDVPYVNMVAGIVLSALAPVHLGPPSSRMAL----- 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 71 VGLLNGLVILITYYTIFKRLKTMPTDTYLLANLAVADILFLILLP-----MAYSEAKS 121

QY 95 WSGGEVHVAAIQIPNNSSLVAMTSTALLSDHITIERALPRTMASVYNTRH----- 146
   : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 WIFGYLCKGFGFYKYSFSGMILLICISIDRYV-----AIQVAVSRHRRARVLLIS 175

QY 147 --VCGFVWGAHLNFSLSLF-----YICSHVSTRALCACQNMLEADATLVF 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 KLSGCGIMMLLFPISIPELLYSGLOKNSGEDTLRCSLVSQAQVALLITIQVAQ-----MVF 230

QY 194 IGYVVPALATLYALVLSRVRRDEPTPLDRDGRLEPSAHRLLVAVTCQFGLMTPHYLIL 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 -GFLVPLMLAMSPCYLIIRITLLQARNPERN-----KAIKYIIAVVVVFIVFQLPYNGV 283

QY 254 LGHTVY---ISRGKRPVDHAYVGLLHFPVDFSKLLAFSSSPVTPLLIYRYMNSPFSKLQRL 310
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 LAQTVANENITNSCETSQKQINAV---DVTYSLASVRCVNPFLYAFIIVKFKFSDFKL 340

QY 311 MKKLPJC 316
Db 341 FKDLGC 346

```

RESULT 10
B55735
lymphocyte-specific G protein-coupled receptor EB11 - human
N/Alternate names: Burkitt's lymphoma receptor 2; Epstein-Barr virus induced protein 1
C/Species: Homo sapiens (man)
C/Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C/Accession: B55735; S52443
R/Schweikardt, V.L.; Raport, C.J.; Godliska, R.; Byers, M.G.; Eddy Jr., R.L.; Shows, T.B.
Genomic 23, 643-650, 1994
A/Title: Cloning of human and mouse EB11, a lymphoid-specific G-protein-coupled receptor
A/Reference number: A55735; MUID:95154835; PMID:7851893
A/Accession: B55735
A/Status: Preliminary
A/Molecule type: mRNA
A/Residues: 1-378 <SCH>
A/Cross-references: KEMIPOT:P32248; GB:L31581; NID:g468319; PIDN:AAA74231.1; PID:g468320
R/Burgstahler, R.; Kempkes, B.; Staedle, K.; Lipp, M.
Submitted to the EMBL Data Library, February 1995
A/Description: The expression of the chemokine receptor BUR2/EB11 is specifically transcribed
A/Reference number: S52443
A/Accession: S52443
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 21-378 <BUR>
A/Cross-references: EMBL:X84702
C/Genetic8:
A/Gene: GDB:CMKBR7; EB11; BUR2; CCR7
A/Cross-references: GDB:342065; OMIM:600242
A/Map position: 17q12-17q21.2
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor

Query Match	10.2%;	Score 177.5;	DB 2;	Length 378;
Best Local Similarity	23.9%;	Pred. No. 3.6e-07;		
Matches 76;	Conservative 49;	Mismatches 146;	Indels 47;	Gaps 10;

```

RESULT 11
S26667
G protein-coupled receptor BLR1 - human
C/Species: Homo sapiens (man)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S26667
R/Dobner, T.; Wolf, I.; Emrich, T.; Lipp, M.
Eur. J. Immunol. 22, 2795-2799, 1992
A/Title: Differentiation-specific expression of a novel G protein-coupled receptor from
A/Reference number: S26667; MUID:93049615; PMID:1425907
A/Accession: S26667
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-372 >DOB>
A/Cross-references: UNIROT:P32302; EMBL:X60149; NID:g29459; PIDN:CAA48252.1; PID:g2946
C/Genetics:
A/Gene: GDB:BLR1
A/Cross-references: GDB:136235; OMIM:601613
A/Map position: 15q26.1-15q26.1
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match      10.2%; Score 177; DB 2; Length 372;
Best Local Similarity 23.6%; Pred. No. 3.9e-07;
Matches 84; Conservative 60; Mismatches 126; Indels 86; Gaps 17;

      7  FNGTGLVEE-LPPACD-----LOGGLSLSLGLGVGVPGVGLCYNALVLAH 54
      27 YNDTSLVENHLCPRTSEPPLMASFKAVFVPAYSILFLLGVGNLV-----LVILERH 79
QY 55 SKASMTWPDVYFVNMAVAGLVLSALAEVHLGPSSRMALMSVGEVHVAL-QIPFNVS 113
Db 27 YNDTSLVENHLCPRTSEPPLMASFKAVFVPAYSILFLLGVGNLV-----LVILERH 79
QY 80 -RGRSSTETFFLLHVAADLLVETLP-FAVAEGSVGVMLGTFLPKTVIALHKNFYCSS 137
Db 80 -RGRSSTETFFLLHVAADLLVETLP-FAVAEGSVGVMLGTFLPKTVIALHKNFYCSS 137
QY 114 LVAMYSFALLSLDHYIRALPRTYMASVYTRH-----VGFVWGGALLTSPSSLLF 165
Db 114 LVAMYSFALLSLDHYIRALPRTYMASVYTRH-----VGFVWGGALLTSPSSLLF 165
QY 138 LL-----LACTAVDRYL-----AIVHAVVHYRRHRLLSIHITGCTIWLVGFLALPELF 187
Db 138 LL-----LACTAVDRYL-----AIVHAVVHYRRHRLLSIHITGCTIWLVGFLALPELF 187
QY 166 YICS--HVSRALECAKQNAEA---ADATLVFI---GVVPALATLVYALV-LLSRVVR 215
Db 166 YICS--HVSRALECAKQNAEA---ADATLVFI---GVVPALATLVYALV-LLSRVVR 215
QY 188 AKYSQGHNNHNSLRCTFSQENQAEHAWFTSRFLYHAGFLPRLNMGWCYGVGVHRLRQ 247
Db 188 AKYSQGHNNHNSLRCTFSQENQAEHAWFTSRFLYHAGFLPRLNMGWCYGVGVHRLRQ 247
QY 216 EDTPLDRDTRGLRPSAHRLLVATVCTQPGMLTPHYLLLGHTVILISRGKPD----- 267
Db 216 EDTPLDRDTRGLRPSAHRLLVATVCTQPGMLTPHYLLLGHTVILISRGKPD----- 267

```

Db 248 AQRPRQ-----KAKVALVTSIFLCSFPHIVFLDT--LAKKAVDNTCKINGS 299
QY 268 -----AHYGLHFVKDFSKLLAFSSFTVPLLYRYNQSPPSKLQRLMKKLP 316
Db 300 LPVAITWCFGLAH-----CCINPMLTYFAGVKFRSDLSRLTLKGC 342

RESULT 12

JE0349
Interferon-inducible protein 10 (IP-10) receptor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C/Accession: JE0349
R/Tamari, M.; Tomihaga, Y.; Yatesman, K.; Narumi, S.
Biochem. Biophys. Res. Commun. 251, 41-48, 1998
A/Title: Cloning of the murine interferon-inducible protein 10 (IP-10) receptor and its
A/Reference number: JE0349; MUID:9909219; PMID:9790904
A/Accession: JE0349
A/Molecule type: mRNA
A/Residues: 1-367 <TAM>
A/Cross-references: UNIPROT:Q9QW6; DDBJ:AB003174; NID:93798731; PIDN:BA034045.1; PID:93
C/Comment: This protein is important for lymphocyte trafficking to lymphoid organs.
C/Superfamily: Vertebrate rhodopsin

Query Match 10.1%; Score 175.5; DB 2; Length 367;
Best Local Similarity 22.1%; Pred. No. 5,1e-07;
Matches 77; Conservative 59; Mismatches 137; Indels 75; Gaps 14;

QY 8 NGGLVLEELPACODLQGL-----SLSLGLVGVPGVGLCYNALVLNLSKASMTM 61
Db 32 NESDFSPSPCPDPFSINFDRTFLPALYSLLFL-----IGLNGVAVALLSQRRLAS 86
QY 62 PDVYVMAVAGVLSALPVLHLLGPPSSKWL-----MSVGEVVALQIPNVSLVA 116
Db 87 TDFLLHLAVADVLLVTLPL-----MAVDAVQVWFGPLCKVAGALENINFYAG 137
QY 117 MYSTALLSDHY-----IERALPRTYMASVYTRHVGCFWAGALITSPSSLLFYIC 168
Db 138 AFLACISFDRYLSIYHATQIYRDRVAVL-----TCIYWGICLIPALPDL-YS 190
QY 169 SHVSTR--ALECKAKMNAEADATLVFI-----GYVPALATLY-----ALVLSRVVR 215
Db 191 ANYDQRLNTHC--QYNFPQVGTALRVQLVAGFLLPLLVMAVCYAHILAVLSR--- 245
QY 216 EDTPLDRDTGRLEPSAHLRLVATVCTQFGLMTPHYLLIGHYV-----IISRKPVDATL 271
Db 246 -----GQRFRAKRLVVVVVAAFAVCWTPYHLVAVLDIMDVGLARRCGRESH-- 294
QY 272 GLLHFVKDFSKLLAFSSFTVPLLYRYNQSPPSKLQRLMKKLP 319
Db 295 --VDVAKSVTSGWGMHCINPLLYAFVGVKFRQGMWMLFTRLGRSDQ 340

RESULT 13

S29480
bombesin receptor - guinea pig
C/Species: Cavia porcellus (guinea pig)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S29480
R/Gorbulev, V.; Akhundova, A.; Buechner, H.; Fahrenholz, F.
submitted to the EMBL Data Library, July 1992
A/Description: Molecular cloning of a new bombesin receptor subtype.
A/Reference number: S29480
A/Accession: S29480
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-399 <GOR>
A/Cross-references: UNIPROT:P35371; EMBL:X67126; NID:949545; PIDN:CAA47605.1; PID:949546
C/Superfamily: endothelin receptor B
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 9.8%; Score 169; DB 2; Length 399;

Best Local Similarity 22.7%; Pred. No. 1.9e-06;
Matches 68; Conservative 69; Mismatches 119; Indels 44; Gaps 14;

QY 38 VPVGLCYNALVLNLSKASMTMPDVYFVNMAVAG-LVLSALPVLHLLGPPSSKWLMS 96
Db 58 ISVGLGNAILIKVFPKYSMQVTPNITFTSLADQLLLTCVPDATHYLAEGMLFGR 117
QY 97 VGEVVALQIPNVSSLVAMYSTALLSDHY--IERALPRTYMASVYTRHVGCFWVG 154
Db 118 IGCKVLSFRL--TSVGSVFTLITLSADRKAVKAVKPLERPSNAILTKKACICIM 174
QY 155 ALLTSFSSLLFYICHSVST-----RALECA-----KMNAEADATLVFTGYVP 199
Db 175 SMIFALPEAIF--SNVHTLRDPNKNMTSEWCAPYVSEKLOEITHALLSFLVF--YIIP 229
QY 200 --ALATVVALVLSRVEDT---PLDRDT-GRLEPSAHLRLVATVCTQFGV---WTPH 249
Db 230 LSLISVYSLI--ATLVKSTINITPEQSARKKVESRKRIAKTVLVVLFALCWLPN 287
QY 250 YIILGHTVVISRGPVDAHYIGLHF-VKDFSKLLAFSSFTVPLLYRYNQSPPSKLQ 308
Db 288 HLLNLVHSPTRKAYBDSA-----IHFIYVTFSRVLAFSNSCVNPFALYWLKTPQKQFX 342

RESULT 14

S56162
MDCK15 protein - human
C/Species: Homo sapiens (man)
C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
C/Accession: S56162
R/Bareille, L.; Loescher, M.; Tobler, A.; Baggiolini, M.; Moser, B.
Biochem. J. 309, 773-779, 1995
A/Title: Sequence variation of a novel heptahelical leucocyte receptor through alternat
A/Reference number: S56162; MUID:95366951; PMID:7639692
A/Accession: S56162
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-327 <BAR>
A/Cross-references: EMBL:X68829; NID:9840783; PIDN:CAA48723.1; PID:9840784
C/Superfamily: vertebrate rhodopsin

Query Match 9.6%; Score 167; DB 2; Length 327;
Best Local Similarity 23.8%; Pred. No. 2.3e-06;
Matches 77; Conservative 56; Mismatches 117; Indels 74; Gaps 15;

QY 27 SLSLGLVGVPGVGLCYNALVLNLSKASMTMPDVYFVNMAVAGVLSALPVLHLLG 86
Db 14 SLIFLGVIGNVLV-----LVILERH-RQTRSTETPLFLHADVADLLVFLIP-PAVA 64
QY 87 PSSRMLMSVGEVVAL-QIPNVSSLVAMYSTALLSDHYIERALPRTYMASVYNT 145
Db 65 EGSVGMVLGTFLLCTVIALHKNVFCSSL-----LACIADVRLY-----AIYHVAAYR 114
QY 146 H-----VCGFWGALITSPSSLLFYICS--HVSTRALCAKMONAEA--ADATLV 192
Db 115 HRLSLIHICGTIMLVGFLALPELFAVSGCHNNNSLPCTFSQENADETHAWTSR 174
QY 193 FI-----GYVPALATLYALV-LISRVREDTPLDRDTGRLEPSAHLRLVATVCTQFG 247
Db 175 FLVHVAQFLPMLVMGVCYGVVHRQARRPQRQ-----KAVRAILVTSIFLFCWS 228
QY 248 PHYLLIGHVVISRGPVD-----AHYGLHFVNDFSKLLAFSSFTV 292
Db 229 PTHIVILDT--LARKAVDNTCKLNGSLPVAITWCFGLAH-----CCIN 273
QY 293 PLLYRYNQSPPSKLQRLMKKLP 316
Db 274 PMLTYFAGVKFRSDLSRLTLKGC 297

RESULT 15

158186
probable G protein-coupled receptor - rat

C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
 C:Accession: I58186
 R:Harriison, J.K.; Barber, C.M.; Lynch, K.R.
 Neurosci. Lett. 169, 85-89, 1994
 A:Title: cDNA cloning of a G-protein-coupled receptor expressed in rat spinal cord and b
 A:Reference number: I58186; MUID:9432113; PMID:8047298
 A:Accession: I58186
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-354 <RES>
 A:Cross-references: UNIPROT:P35411; EMBL:U04808; NID:G2558635; PIDN:AAB87093.1; PID:9439
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor

Query Match 9.64; Score 167; DB 2; Length 354;
 Best Local Similarity 22.94; Pred. No. 2.5e-06;
 Matches 73; Conservative 54; Mismatches 120; Indels 72; Gaps 14;

```

QY 41 GLCYNLLVLANHSKSMTPDYFVNMVAGLVLSALAPVHLLGPPSSRKALMSVGE 100
Db 46 GLVGNLLVVALTNSRKSISITDIYLLNLALSDLLFVATLPF-----WTHYLLSH 96
QY 101 -VVAL---QIPNVASLVAMYSTALLSLDHYIERALPTYMASVYNTRHV-----C 148
Db 97 GLNNMCKLTATAPFFIGFPGGIFPIYISIDRYLAIVL---AANSMMNTVQHGVTISL 152
QY 149 GPVWGALLTSFSSLLFYICSHVSTRALC-----AKQNAEADATLVFIQY 196
Db 153 G-VMAAILVAAPQFMP-----TKRKDNECIGDYPEVLQEIWPVLRNSE-----VNILGF 201
QY 197 VVPALATLVALLVLSRVRREDTPLDRDTGRLEPSAHLVAVYCTQFGLTTPHYLLDGH 256
Db 202 VLPLLLMSFCYFRIVRT-----LPSCNKRKKAIRLLILVVVFLLFWTPYNIIVLE 255
QY 257 TV-----IISRGKPVDAHYLGILHFPVDPSKLLAFSSFTPLLYRYMNQSPSKLQRLM 311
Db 256 TLKFTYNFPSCGMKRD-----LRMALSVETETVAFSHCCINPFIYAFAGEKFRIRYLAHLY 309
QY 312 KK--LPCGDRHCSPDHMG 327
Db 310 NKCLAVLCG---RPVHAG 324

```

Search completed: December 14, 2004, 20:25:44
 Job time : 15.5627 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: December 14, 2004, 18:37:44 ; Search time 54.3673 Seconds
(without alignments)
3524.170 Million cell updates/sec

Title: US-09-995-225B-2
Perfect score: 1733
Sequence: 1 MMSCSFNGTGLVEBLPACO.....LPGCDHRCSPDHMGVQVTLA 333

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0 %
Maximum Match 100 %
Listing first 45 summaries

Database : uniprot_02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1733	100.0	333	096CH1	096CH1 homo sapien
2	1687	97.3	333	08HXR4	08HXR4 macaca fasc
3	1277	73.7	333	08BHR6	08BHR6 mus musculu
4	1276	73.6	333	099LE2	099LE2 mus musculu
5	1272	73.4	333	08BHU4	08BHU4 mus musculu
6	1092	63.0	333	06PTG9	06PTG9 xenopus lae
7	1092	63.0	333	06PTG9	06PTG9 xenopus lae
8	772	44.5	151	086SP5	086SP5 homo sapien
9	267	15.4	375	06FHV6	06FHV6 homo sapien
10	267	15.4	375	06FHV6	06FHV6 homo sapien
11	233.5	13.5	375	06FHV6	06FHV6 homo sapien
12	229.5	13.2	375	06FHV6	06FHV6 homo sapien
13	223.5	12.9	375	06FHV6	06FHV6 homo sapien
14	208.5	12.0	355	08HZN5	08HZN5 mus musculu
15	206.5	11.9	355	08HZN5	08HZN5 mus musculu
16	205	11.7	358	08HZN5	08HZN5 mus musculu
17	203.5	11.8	358	08HZN5	08HZN5 mus musculu
18	201.5	11.6	353	08HZN5	08HZN5 mus musculu
19	201	11.6	353	08HZN5	08HZN5 mus musculu
20	201	11.6	353	08HZN5	08HZN5 mus musculu
21	200	11.5	353	08HZN5	08HZN5 mus musculu
22	200	11.5	353	08HZN5	08HZN5 mus musculu
23	200	11.5	353	08HZN5	08HZN5 mus musculu
24	198.5	11.5	350	08HZN5	08HZN5 mus musculu
25	198	11.4	350	08HZN5	08HZN5 mus musculu
26	195.5	11.3	350	08HZN5	08HZN5 mus musculu
27	195.5	11.3	350	08HZN5	08HZN5 mus musculu
28	193.5	11.2	350	08HZN5	08HZN5 mus musculu
29	193.5	11.2	350	08HZN5	08HZN5 mus musculu
30	192.5	11.1	350	08HZN5	08HZN5 mus musculu
31	191	11.0	350	08HZN5	08HZN5 mus musculu

32	190	11.0	366	2	0867B2	0867B2 capra hircu
33	188	10.8	366	2	0867B2	0867B2 capra hircu
34	185	10.7	349	1	IL8A_RAT	IL8A_RAT
35	185	10.7	349	1	IL8A_RAT	IL8A_RAT
36	184.5	10.6	374	1	0861S1	0861S1 sus scrofa
37	183	10.6	380	2	BAD06309	BAD06309 sus scrofa
38	183	10.6	380	2	BAD06309	BAD06309 sus scrofa
39	181.5	10.5	378	2	06U2D6	06U2D6 ratu
40	181.5	10.5	378	2	06U2D6	06U2D6 ratu
41	180.5	10.4	359	1	09PYV7	09PYV7 anguilla an
42	180.5	10.4	359	1	09PYV7	09PYV7 anguilla an
43	179.5	10.3	378	2	08CAS2	08CAS2 mus musculu
44	179	10.3	351	2	0810W6	0810W6 mus musculu
45	179	10.3	351	2	AA084305	AA084305 mus musculu

ALIGNMENTS

RESULT 1	ID	Q96CH1	PRELIMINARY;	PRT;	333 AA.
DT	01-DEC-2001	(TREMBlrel. 19, Created)			
DT	01-DEC-2001	(TREMBlrel. 19, Last sequence update)			
DT	01-OCT-2003	(TREMBlrel. 25, Last annotation update)			
DE	G protein-coupled receptor 146.				
GN	Name=GPR146;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISUS=Kidney;				
RX	MDL=22388257; PubMed=12477932;				
RA	Strausberg R.L., Pelngold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,				
RA	Diatcenco L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Shapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Baba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,				
RA	Villalton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,				
RA	Krzywnicki M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,				
RA	Jones S.J., Marra M.A.,				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISUS=Kidney;				
RA	Strausberg R.,				
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC014241; AA014241.1; -.				
DR	Genew; HNC:21718; GPR146.				
DR	GO; GO:0016021; C:integral to membrane; IEA.				
DR	GO; GO:0004872; F:receptor activity; IEA.				
DR	GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.				
DR	GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.				
DR	InterPro; IPR000276; GPCR_Rhodopsin.				
DR	Pfam; PF00001; 7tm_1; 1.				
DR	PRINTS; PR00237; GPCRHHODPSN.				
DR	PROSITE; PSS0262; G_PROTEIN_RECPT_FL_2; 1.				
KW	Receptor.				
SQ	SEQUENCE 333 AA; 36580 MW; BB8D6505A6C2D9F4 CRC64;				

Query Match 100.0%; Score 1733; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.6e-122;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMSCGWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANLHRSKASMT 60
DB 1 MMSCGWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANLHRSKASMT 60
QY 61 MPDYVFNNAVAGLVLSALAPVHLLGPSSRMALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFNNAVAGLVLSALAPVHLLGPSSRMALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYIERALPRTYMASVYNTNRHVCGFWMGALLTSPSSLLFTYICSHVSTRALAECAK 180
DB 121 ALLSDHYIERALPRTYMASVYNTNRHVCGFWMGALLTSPSSLLFTYICSHVSTRALAECAK 180
QY 181 MONEAADATLVFITYGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONEAADATLVFITYGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPLLYRYNM 300
DB 241 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPLLYRYNM 300
QY 301 QSPFSKLRMLMKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSPFSKLRMLMKLPCGDRHCSPDHMGVOQVLA 333

RESULT 2
Q8HXF4 PRELIMINARY; PRT; 333 AA.
ID 08HXF4
AC 08HXF4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypoetical protein.
DE Hypoetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
NCBI_TaxID=9541;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RX MEDLINE=21458551; PubMed=11574149;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirata M., Suto Y.,
RA Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K.,
RT "Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
RT chromosomes.";
RL Gene 275:31-37(2001).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Sugano S.,
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB093636; BAC21610.1; -;
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO; GO:007186; P: G-protein coupled receptor protein signaling. . .; IEA.
DR InterPro; IPR00237; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00262; G_PROTEIN_RECPT_F1_2; 1.
KM Hypoetical protein.
SQ SEQUENCE 333 AA; 36365 MW; 9D95B0461F93439 CRC64;

Query Match 97.3%; Score 1687; DB 2; Length 333;
Best Local Similarity 96.4%; Pred. No. 4.6e-119;
Matches 321; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 MMSCGWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANLHRSKASMT 60

DB 1 MMSCGWNGTGLVEELPACODLQGLSLSLGLVGVPGVGLCYNALVLVLANLHRSKASMT 60
QY 61 MPDYVFNNAVAGLVLSALAPVHLLGPSSRMALMSVGEVHVALQIPNVSSLVAMYST 120
DB 61 MPDYVFNNAVAGLVLSALAPVHLLGPSSRMALMSVGEVHVALQIPNVSSLVAMYST 120
QY 121 ALLSDHYIERALPRTYMASVYNTNRHVCGFWMGALLTSPSSLLFTYICSHVSTRALAECAK 180
DB 121 ALLSDHYIERALPRTYMASVYNTNRHVCGFWMGALLTSPSSLLFTYICSHVSTRALAECAK 180
QY 181 MONEAADATLVFITYGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
DB 181 MONEAADATLVFITYGVVPALATLYALVLSRVREDTPLDRDTGRLEPSAHLVAATVC 240
QY 241 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPLLYRYNM 300
DB 241 TQFGLMTPHYLLILGHTVLIISRGKPYDAHYLGILHFVKDPSKLLAFSSSFVTPLLYRYNM 300
QY 301 QSPFSKLRMLMKLPCGDRHCSPDHMGVOQVLA 333
DB 301 QSPFSKLRMLMKLPCGDRHCSPDHMGVOQVLA 333

RESULT 3
Q8BHR6 PRELIMINARY; PRT; 333 AA.
ID 08BHR6
AC 08BHR6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
DE enriched library, clone: B230399N13 product: HYPOHETICAL 36.6 kDa
DE PROTEIN full insert sequence.
GN Name=B230399N13;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN 14
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20499574; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN 15
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;

RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Iehli Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujitake S., Inoue K., Togawa K., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-formet
 sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RP [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirakawa T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurahara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi P., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK046512; BAC32762.1; -
 DR MGD; MGI:1933113; BC003323.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro: IPR000226; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE; PRO0237; GPCRHRDOPSN.
 DR PRINTS; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 DR KX Hypothetical protein.
 SQ SEQUENCE 333 AA; 36571 MW; C51690838AB42988 CRC64;

Query Match 73.7%; Score 1277; DB 2; Length 333;
 Best Local Similarity 74.5%; Pred. No. 3.6e-88;
 Matches 248; Conservative 31; Mismatches 52; Indels 2; Gaps 2;

QY 1 MMSGSPFNGTGVYEBLPAQCDLQGLSLSLGLVVGVPVGLCYNALLVLANHSKASMT 60
 DB 1 MMSGSPFNGTGVYEBLPAQCDLQGLSLSLGLVVGVPVGLCYNALLVLANHSKASMT 59
 QY 61 MPDVYFVNNAVAGLVSLAPVHLGLPSSRWALMSVGSVHVALQIPFNVSLLVAMYST 120
 DB 60 MPDVYFVNNAVAGLVSLAPVHLGLPSSRWALMSVGSVHVALQIPFNVSLLVAMYST 119
 QY 121 ALISLDHYIERALPRTYMASVYTRHYVCGFWGAGALITFSLSLLFYICSHVSTRALECAK 180
 DB 120 ALISLDHYIERALPRTYMASVYTRHYVCGFWGAGALITFSLSLLFYICSHVSTRALECAK 179
 QY 181 MONAABADTLVIGVVPPLATLVLTALVLTLSRVRRBETPLDRDTGSLPEAHRLVLAATVC 240
 DB 180 MONAABADTLVIGVVPPLATLVLTALVLTLSRVRRBETPLDRDTGSLPEAHRLVLAATVC 239
 QY 241 TQFGLTTPHYLLGLHTVLIISRGKPYDAHYTLGLHFVKQPSKILAFSSSFVTPLVRYNM 300
 DB 240 TQFGLTTPHYLLGLHTVLIISRGKPYDAHYTLGLHFVKQPSKILAFSSSFVTPLVRYNM 298
 QY 301 QSPFSTLQRLMKGLPGGDRHCSPDHMGVQVTLA 333
 DB 299 KAFPGKLRILMKMKHCGRRHCSPPSGIQVMA 331

RESULT 4
 ID 099LE2 PRELIMINARY; PRT; 333 AA.
 AC 099LE2;
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE CDNA sequence BC003333 (Mus musculus NOD-derived CD1c +ve dendritic
 DE cell cDNA, RIKEN full-length enriched library, clone:R630050602
 DE product:HYPOPHYSIN 36.6 kDa PROTEIN, full insert sequence).
 GN Name=BC003333;
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feilgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebli T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallory S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzyzaniak M.I., Skalska U., Smalins D.E., Scherach A., Schein J.E.,
 RA Jones S.J., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN PANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RA The PANTOM Consortium;
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team:
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD;
 RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1157-1771(2000).
 [8]

RP SEQUENCE FROM N.A.

RC STRAIN=MOD;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK089323; BAC40804.1; -;
 DR MGD; MGI:193313; BC003323.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007166; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm1.1;
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR Hypothetical protein.
 KY SEQUENCE 333 AA; 36557 MW; 5F76909350D801F4 CRC64;
 SQ

Query Match 73.6%; Score 1276; DB 2; Length 333;
 Best Local Similarity 74.5%; Pred. No. 4.3e-88;
 Matches 248; Conservative 31; Mismatches 52; Indels 2; Gaps 2;

QY 1 MMSGFWNGTGLVEBELPACODLQGLSLSLGLVVGVGVCYNALLVTLALHLSKASMT 60
 DB 1 MMSGGLPLNSTAMAE-PLCNRLRLGLMTVLSLYLGCGVSGYNALLVTLALHLSKASMT 59
 QY 61 MPVYFVNMAVAGLVLSALAPHLCPSSRNALMSVGGEVVALQIPFNVSILVAMYST 120
 DB 60 MPVYFVNMAVAGLVLTALAPVLLGPASRNALMSISEAHVTLILFNVASLVMTYST 119
 QY 121 ALLSDHYTERALPRTYMASVYNTRVGFWGVALTSFSSLLTYTCSHYSTRALBCAK 180
 DB 120 ALLSDHYTERALPRTYMASVYNTRVGFWGVALTSFSSLLTYTCSHYSTRALBCAK 179
 QY 181 MONAARADATVFIGVVPALATLVALLSFRREDTPDLDRDTGRLEPSSAHLRLVATVC 240
 DB 180 MONAARADATVFIGVVPALATLVALLSFRREDTPDLDRDTGRLEPSSAHLRLVATVC 239
 QY 241 TOFGMLTPHYLLIGHTVVISRKGVDVDAHYLGHLFVDFSKLAFSSFTPLLYRYM 300
 DB 240 TOFGMLTPHYLLIGHTVVISRKGVDVDAHYLGHLFVDFSKLAFSSFTPLLYRYM 298
 QY 301 OSFPEKLOLMKGLPCGDRHSCPDHMGVOQVLA 333
 DB 299 KAFPGKRLRMKKMKCGRRHSCPDPSGQVMA 331

RESULT 5
 OBBH04 PRELIMINARY; PRT; 333 AA.
 AC OBBH04
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
 DE library, clone:R030026M12 product:HYPOTHEICAL 36.6 kDa PROTEIN, full
 DE insert sequence.
 GN Name=BC003323;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA The FANTOM Consortium;
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT 60,770 full-length mouse transcriptome based on functional annotation of
 RT Nature 420:563-573(2002).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1157-1771(2000).
 [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK087100; BAC39803.1; -;
 DR MGD; MGI:193313; BC003323.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.

DR GO:0007186; P:G-protein coupled receptor protein signalin. . . IEA.
DR Interpro: IPR000276; GPCR_Rhodopn.

DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.

DR PROSITE: PSS0262; G_PROTEIN_RECIP_F1_2; 1.
KW Hypothetical protein.

SEQUENCE 333 AA; 36597 MW; C5168D28B71F5988 CRC64;

Query Match 73.4%; Score 1272; DB 2; Length 333;
Best Local Similarity 74.2%; Pred. No. 8.5e-88;
Matches 247; Conservative 31; Mismatches 53; Indels 2; Gaps 2;

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVLCYNALLVLANHLSKASMT 60
D 1 MMSCSGLNSTANMAEE-PLCNRLRGLVNLISLYGAGVPSLSGNALLVLANHLSKASMT 59
QY 61 MPDVFYVMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSLSVANYST 120
D 60 MPDVFYVMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSLSVANYST 119
QY 121 ALLSLDHYIERALPRTMASVYNTNRYHCGFVWGALITSSSLFYICSHVSTALBCKAK 180
D 120 ALLSLDHYIERALPRTMASVYNTNRYHCGFVWGALITSSSLFYICSHVSTALBCKAK 179
QY 181 MONAADAATLVFIQYVVPALATLYALVLSRVREPTPLDRDGTGLEPSAHLRLVATVC 240
D 180 MONAADAATLVFIQYVVPALATLYALVLSRVREPTPLDRDGTGLEPSAHLRLVATVC 239
QY 241 TOFGLMTPHYLLILGHTVVISRGPVDAHYLGHLFVKDPSKLAFASSFTPLLYRYNM 300
D 240 TOFGLMTPHYLLILGHTVVISRGPVDAHYLGHLFVKDPSKLAFASSFTPLLYRYNM 298
QY 301 QSFPSKLGRLMKLPCGDRHCSPPHMGVQVLA 333
D 299 KAFPGKLRMLKMKHCGRRHCSPPDGGIQVMA 331

RESULT 6
Q6P7G9 PRELIMINARY; PRT; 333 AA.
AC Q6P7G9; (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE MCG68817 protein.
GN Name=MCG68817;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OC Xenopodidae; Xenopus.
ON NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Datchenko L., Marusik A., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Tishiyki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wozley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kesteman M., Madan A.C., Schechenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Kravitz M.I., Skalek U., Small D.E., Schnerch A., Schein J.B.,
RA Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RC TISSUE=Lung;
RX MEDLINE=22341132; PubMed=12454917;

RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative";
RL Dev. Dyn. 225:384-391 (2002).

QY 1 MMSCSWNGTGLVEBELPACODLQGLSLSLGLVGVPGVLCYNALLVLANHLSKASMT 60
D 1 MMSCEDLNTNNSGEQYLCNEFHLFLIFSVLLIICFPVGLCYNQVLVNLNKAHMT 60
QY 61 MPDVFYVMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSLSVANYST 120
D 61 MPDVFYVMAVAGLVLSALAPVHLGPPSSRMALMSVGEVHALQIPFNVSLSVANYST 120
QY 121 ALLSLDHYIERALPRTMASVYNTNRYHCGFVWGALITSSSLFYICSHVSTALBCKAK 180
D 121 TLISLDYIECALPRTMSSVYNTNRYHCGFVWGALITSSSLFYICSHVSTALBCKAK 180
QY 181 MONAADAATLVFIQYVVPALATLYALVLSRVREPTPLDRDGTGLEPSAHLRLVATVC 240
D 181 MONAADAATLVFIQYVVPALATLYALVLSRVREPTPLDRDGTGLEPSAHLRLVATVC 240
QY 241 TOFGLMTPHYLLILGHTVVISRGPVDAHYLGHLFVKDPSKLAFASSFTPLLYRYNM 300
D 241 TOFGLMTPHYLLILGHTVVISRGPVDAHYLGHLFVKDPSKLAFASSFTPLLYRYNM 300
QY 301 QSFPSKLGRLMKLPCGDRHCSPPHMGVQVLA 332
D 301 KNFSGKLRMLKMKHCGRRHCSPPDGGIQVMA 332

RESULT 7
AAH61674 PRELIMINARY; PRT; 333 AA.
AC AAH61674;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE MCG68817 protein.
GN Name=MCG68817;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OC Xenopodidae; Xenopus.
ON NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT Initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,
Altehusen S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang Y., Hsieh F.,
Diachenko L., Marusha K., Farmer A.A., Rubin G.M., Hong R.,
Stapleton M.J., Soares M.B., Bonaldi M.F., Casavant T.J., Scheetz T.E.,
Bromstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
Vallalun D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Klein S., Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC061674; AAH61674.1; -
SQ SEQUENCE 333 AA; 38007 MW; EB72DC50EC8EB815 CRC64;

Query Match 63.0%; Score 1092; DB 2; Length 333;
Best Local Similarity 60.8%; Pred. No. 3.1e-74;
Matches 202; Conservative 61; Mismatches 69; Indels 0; Gaps 0;

QY 1 MMSGFWNGTGLVEELPACODLQGLSLGLGVVGVPGLCYNMLLVIANHGRKASMT 60
DB 1 MMSCEDLNTNNGEEQYLNCNEPFLFPIFSVLYLIICFPGGLCYNQVLVLYNKAWT 60
QY 61 MPDYVFNMAVAGLVSLAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYS 120
DB 61 MPDYVFNMAVAGLVSLAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYS 120
QY 121 ALLISDHYIERALPRTYMASVYNTRHVCGFWGALLTFSFSLFYICSHVSTRALCEA 180
DB 121 TLLSLDHYIECALPRTYMSVYNTKHCPIWCGAVLTSFSSILFYICNHVSFKIECSK 180
QY 181 MKNAEADATLVPIGTVVPPALATLVALLSVRRREDTLPDRGTGLSPAHGLVAVYC 240
DB 181 MKNREAADIMVLIGVVPPIAVLAVLILQIRKATPDDQSGRLDPSVHRLIATYC 240
QY 241 TOFGLMTPHYLILGHTVLIISRGKPVDAHYLGHLHFQDPSKLLAFSSSFVTPLRYNM 300
DB 241 TOPIIMTPPYVTLVNTVPMFARVKSNTFYIRIFQTEGSLNLAFLSSSVPLIRHIN 300
QY 301 QSFPSKLQRLMKLPCGDRHCSPDHNGVQVL 332
DB 301 KNFSGLQRLKRLKRLHCGSGCGCTHEHTVQOV 332

RESULT 8
Q86SP5 PRELIMINARY; PRT; 151 AA.
AC Q86SP5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE G protein-coupled receptor PGR8 (Fragment).
OS Homo sapiens (Human).
RA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22584407; PubMed=12679517;
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
Bergman J.E., Galtanaris G.A.;
RT "The G protein-coupled receptor repertoires of human and mouse.";
Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
RL EMBL: AY255535; AA085047.1; -
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0001861; F:G-protein coupled receptor protein signalin. . .; IEA.
DR Interpro: IPR000276; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 151
FT NON_TER 151 151
SQ SEQUENCE 151 AA; 16325 MW; 42FE8329C47C699A CRC64;

Query Match 44.5%; Score 772; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 TMDYVFNMAVAGLVSLAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYS 119
DB 1 TMDYVFNMAVAGLVSLAPVHLGPPSSRWALMSVGEVHVALQIPFNVSILVAMYS 60
QY 120 TLLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTFSFSLFYICSHVSTRALCEA 179
DB 61 TLLSLDHYIERALPRTYMASVYNTRHVCGFWGALLTFSFSLFYICSHVSTRALCEA 120
QY 180 MKNAEADATLVPIGTVVPPALATLVALLSVRRREDTLPDRGTGLSPAHGLVAVYC 210
DB 121 MKNREAADATLVPIGTVVPPALATLVALLSVRRREDTLPDRGTGLSPAHGLVAVYC 151

RESULT 9
CML2 HUMAN STANDARD; PRT; 375 AA.
AC Q99537; O00143; Q43494; Q13631; Q96F42; Q99981;
DT 15-JUN-1998 (Rel. 36, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Chemokine receptor-like 2 (IL8-related receptor DRY12) (flow-induced
endothelial G protein-coupled receptor) (FEG-1) (G protein-coupled
receptor GPR30) (GPCR-BR).
GN Name=GPR30; Synonyms=CMKRL2, DRY12, CEPR;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97079175; PubMed=8920907;
RA Owan C.S.O., Bly P., Nilsson C., Lolait S.J.;
RT "Cloning of human cDNA encoding a novel heptahelix receptor expressed
in Burkitt's lymphoma and widely distributed in brain and peripheral
tissues.";
RL Biochem. Biophys. Res. Commun. 228:285-292(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97286314; PubMed=914181;
RA Kvingedal A.M., Smeland E.B.;
RT "A novel putative G-protein-coupled receptor expressed in lung, heart
and lymphoid tissue.";
RL FEBS Lett. 407:59-62(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97224403; PubMed=9070864;
RA Feng Y., Gregor P.;

RT "Cloning of a novel member of the G protein-coupled receptor family
 RT related to peptide receptors.";
 RT Blochem. Biophys. Res. Commun. 231:651-654(1997).
 RN [4]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RA McCoy R.L., Perlmuter D.H.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98063308; PubMed=9398636;
 RA Takada Y., Kato C., Kondo S., Korenaga R., Ando J.;
 RT "Cloning of cDNAs encoding G protein-coupled receptor expressed in
 RT human endothelial cells exposed to fluid shear stress.";
 RT Blochem. Biophys. Res. Commun. 240:737-741(1997).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035885; PubMed=9367686;
 RA Carmeci C., Thompson D.A., Ring H.Z., Francke U., Weigel R.J.;
 RT "Identification of a gene (GPR30) with homology to the G-protein-
 RT coupled receptor superfamily associated with estrogen receptor
 RT expression in breast cancer.";
 RT Genomics 45:607-617(1997).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98140132; PubMed=9479505;
 RA O'dowd B.P., Nguyen T., Marchese A., Cheng R., Lynch K.R.,
 RA Heng H.H.Q., Kolakowski L.F. Jr., George S.R.;
 RT "Discovery of three novel G-protein-coupled receptor genes.";
 RT Genomics 47:310-313(1998).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Uterus;
 RX MEDLINE=92388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedlin T.B., Tohllyuki S., Carinici P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richardson D.K., Wozley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalloe D.E.,
 RA Scherach A., Schein J.E., Jones S.U.M., Maria M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Orphan receptor, possibly for a chemokine.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in lung, heart, liver, brain and
 CC lymphoid tissue, in brain regions, expressed as a 2.8 kb
 CC transcript in basal forebrain, frontal cortex, thalamus,
 CC hippocampus, caudate and putamen.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; Y08162; CAA69354.1; -;
 CC DR EMBL; X98510; CAA67133.1; -;
 CC EMBL; U77827; AAC51173.1; -;

DR EMBL; U58828; AAB02736.1; -;
 DR EMBL; AF015257; AAC51904.1; -;
 DR EMBL; U63917; AAB88017.1; -;
 DR EMBL; AF027956; AAC52027.1; -;
 DR EMBL; BC01634; AAH1634.1; -;
 DR PIR; G02670; G02670.
 DR PIR; JCS069; JCS069.
 DR Genew; HGNC:4485; GPR30.
 DR MIM; 601805; -;
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0004930; F: G-protein coupled receptor activity; TAS.
 DR GO; GO:0007186; P: G-protein coupled receptor protein signaling. . . ; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR InterPro; IPR002286; P2_purinocptor.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1;
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 FT DOMAIN 1 62
 FT TRANSMEM 63 84
 FT DOMAIN 85 96
 FT TRANSMEM 97 120
 FT DOMAIN 121 132
 FT TRANSMEM 133 153
 FT DOMAIN 154 175
 FT TRANSMEM 176 194
 FT DOMAIN 195 220
 FT TRANSMEM 221 236
 FT DOMAIN 237 259
 FT TRANSMEM 260 280
 FT DOMAIN 281 306
 FT TRANSMEM 307 327
 FT DOMAIN 328 375
 FT DISULFID 130 207
 FT CARBOHYD 25 25
 FT CARBOHYD 32 32
 FT CARBOHYD 44 44
 FT CONFLICT 16 16
 FT CONFLICT 20 21
 FT CONFLICT 32 49
 FT CONFLICT 104 104
 FT CONFLICT 140 146
 FT CONFLICT 149 179
 FT CONFLICT 184 184
 FT CONFLICT 312 312
 FT CONFLICT 320 320
 FT CONFLICT 355 355
 FT CONFLICT 358 358
 SQ SEQUENCE 375 AA; 42247 MW; 0A54EBDA658F075 CRC64;
 Query Match 15.4%; Score 267; DB 1; Length 375;
 Best Local Similarity 27.4%; Pred. No. 5.1e-12;
 Matches 96; Conservative 65; Mismatches 131; Indels 58; Gaps 17;
 QY 8 NGNGVVEHLPACDDLGSLSLGLGVNPPVGLCTNALLVLANLSKASMTMPDVPYV 67
 DB 44 NGTG--ELSEHQYVIGL-FISCLYTFIFLPPGPNGLILVNVISFRKMTIPDLFI 99
 QY 68 NMAVAGVLSALAPVHLGPPSSRMALMSVGEVHVALQIPFNVSLVAMYSTAL----L 123
 DB 100 NLAVADILVADSLIEFNFHRYVDIAVCTMSLFLQ-----VNMYSSVFLPTWM 151
 QY 124 SLDIYERLALPTVMASVNTNR---VCGFWGALLISFSSLLFYTCISVSTRALECA 179
 DB 152 SPDRYI--ALARMRCSLFRTKHAHRLSCGLI--MASVSATLVPTAVHLQHTBACF 206
 QY 180 KQGNAAEAADTVYFIVGVV--ALATIXALVLSRYRREPTLPDRDTCGRLEP---SAHRL 234
 DB 207 CPADVREVQWLEVTGLGIVPAITIGCYSLIVLVPAH----RHNG-LRPQRQALRM 260
 QY 235 LVATVCTQFGIATPHTYILIGHVITISRGK-----VDAYVIGLILFVDFSKLLA 285

```

Db      261 ILAVLVVFPVCMLEPENFVSVH--LLQRTQPGARPCQKQSFRAHPL-TGHIV-----NLAA 313
Qy      286 FSSSFYTPLLRYNMNSPFSKIQRLMK---LPCGDRHCSPDHMGVQVLT 332
Db      314 FNSNCINPLIYSFLGTFPRDKRLTYIEQKTNLPALNRFCC---HAALKAVI 360

RESULT 10
Q6FHU6      PRELIMINARY; PRT; 375 AA.
ID 06FHU6;
AC 06FHU6;
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DE 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE GPR30 protein (Fragment).
GN Name=GPR30;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RA Halleck A., Ebert L., Moundinya M., Schick M., Eisenstein S.,
RA Neubert P., Kstrang K., Schatten R., Shen B., Henze S., Mar W.,
RA Korn B., Zuo D., Hu Y., Labaer J.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
DR EMBL; CR541655; CAG646456.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Receptor; Transmembrane.
FW NON TER 375
SQ SEQUENCE 375 AA; 42194 MW; 5A54FDB8A2E36AE CRC64;

Query Match 15.1%; Score 261; DB 2; Length 375;
Best Local Similarity 26.9%; Pred. No. 1.5e-11;
Matches 94; Conservative 66; Mismatches 132; Indels 58; Gaps 17;

Qy      8 NGTGVBEIPACDQLGSLSLGLVGVPGVGLCYNMLVLVLANIHSKASMTMPDVYFV 67
Db      44 NGTG---ELSEHQQVYIAL-FLSCLYTIPLFPFGVGNILIVVNISFREKXTIPDLYFI 99
Qy      68 NMAVAGVLSALAPVHLGPPSSRMALMSGGEVHALQIPFVNSLVAMYSTAL-----L 123
Db      100 NLAVADLLVADSLIEVFNHERYVIAVLCFTMSLFLD-----VMYSSVFELTWM 151
Qy      124 SLDPHYERALPRTYMASVYNTNH---VCGFWGALITFSLSLFIYICSHVSTRALBCA 179
Db      152 SFPRYI--ALARMRCSLFRTHKHAHLSCGLIW---MSVSATLIVPFAVHLQHTDEACF 206
Qy      180 KQONAAADATLVFVIGVVP--ALATLVAVLSRRRBDTPLDDTGLEPSAHR---L 234
Db      207 CPADVAREVWLEVTLSGFIPEAIIIGCYSLIVRLVRAH-----SHRG-LPRRQKALCM 260
Qy      235 LVAVTCTQGLWTPHYLLLGHTVLIISRGK-----VDAYLGLLHFQKPSKTLA 285
Db      261 ILAVLVVFPVCMLEPENFVSVH--LLQRTQPGARPCQKQSFRAHPL-TGHIV-----NLAA 313
Qy      286 FSSSFYTPLLRYNMNSPFSKIQRLMK---LPCGDRHCSPDHMGVQVLT 332
Db      314 FNSNCINPLIYSFLGTFPRDKRLTYIEQKTNLPALNRFCC---HAALKAVI 360

RESULT 11
CML2_RAT
ID CML2_RAT
AC 008878; STANDARD; PRT; 375 AA.

```

```

DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE Chemokine receptor-like 2 (G-protein coupled receptor GPR41).
GN Name=Cmkrl2; Synonyms=Gpr41;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Lung;
RX MEDLINE=97312546; PubMed=9168987;
RA Bonini J.A., Anderson S.M., Steiner D.F.;
RT "Molecular cloning and tissue expression of a novel orphan G protein-
RT coupled receptor from rat lung."
RL Biochem. Biophys. Res. Commun. 234:190-193(1997).
CC -1- FUNCTION: Orphan receptor; possibly for a chemokine.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Highly expressed in brain and lung, lower
CC expression in heart, skeletal muscle and kidney.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U92802; AAC53208.1; -.
DR FIR; UC5509; UC5509.
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR InterPro; IPR002286; P2_punocptor.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Glycoprotein; Transmembrane.
KW DOMAIN 1
FT TRANSMM 63 84
FT DOMAIN 1 (Potential).
FT TRANSMM 85 96
FT DOMAIN 2 (Potential).
FT TRANSMM 97 120
FT DOMAIN 3 (Potential).
FT TRANSMM 121 132
FT TRANSMM 133 153
FT TRANSMM 154 175
FT TRANSMM 176 194
FT TRANSMM 195 220
FT TRANSMM 221 236
FT TRANSMM 237 259
FT TRANSMM 260 280
FT TRANSMM 281 306
FT TRANSMM 307 327
FT TRANSMM 328 375
FT DOMAIN 4 (Potential).
FT DISULFID 130 207
FT CARBOHYD 32 32
FT CARBOHYD 44 44
FT CARBOHYD 44 44
SQ SEQUENCE 375 AA; 42259 MW; 08B60576BDBE314 CRC64;

Query Match 13.5%; Score 233.5; DB 1; Length 375;
Best Local Similarity 24.6%; Pred. No. 1.7e-09;
Matches 85; Conservative 69; Mismatches 138; Indels 53; Gaps 16;

Qy      12 LVEELPACDQLGSLSLGLVGVPGVGLCYNMLVLVLANIHSKASMTMPDVYFVNAV 71
Db      45 LTGDLSEHQQVYIAL-FLSCLYTIPLFPFGVGNILIVVNISFREKXTIPDLYFINLAA 103
Qy      72 AGVLSALAPVHLGPPSSRMALMSGGEVHALQIPFVNSLVAMYSTAL-----LSLDH 127
Db      104 ADLLVADSLIEVFNLDGEYVIAVLCFTMSLFLD-----NMYSVFELTWMSPDR 155
Qy      128 YIERALPRTYMASVYNTNH---VCGFWGALITFSLSLFIYICSHVSTRALBCAKQON 183

```


DE Mus musculus adult male pituitary gland cDNA, RIKEN full-length
 DE enriched library, clone:533040418 product:CHEMOKINE RECEPTOR-LIKE 2
 DE (G-PROTEIN COUPLED RECEPTOR GPR41) homolog.
 GN Name=Gpr30;
 OS Mus musculus (Mouse);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SOURCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX MEDLINE=99279253; PubMed=1039636;
 RA Carninci P.; Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SOURCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RA the FANTOM Consortium;
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P.; Shibata Y.; Hayatsu N.; Sugahara Y.; Shibata K.; Itoh M.;
 RA Kono H.; Okazaki Y.; Muramatsu M.; Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K.; Itoh M.; Aizawa K.; Nagao S.; Sasaki N.; Carninci P.;
 RA Kono H.; Akiyama U.; Nishi K.; Kitsuunai T.; Tashiro H.; Itoh M.;
 RA Suni N.; Ishii Y.; Nakamura S.; Hazama M.; Nishino T.; Harada A.;
 RA Yamamoto R.; Matsumoto H.; Sakaguchi S.; Ikegami T.; Kashiwagi K.;
 RA Fujiwaka S.; Inoue K.; Togawa Y.; Izawa M.; Ohara E.; Watanishi M.;
 RA Yoneda Y.; Ishikawa T.; Ozawa K.; Tanaka T.; Matsura S.; Kawai U.;
 RA Okazaki Y.; Muramatsu M.; Inoue Y.; Kita A.; Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=pituitary gland;
 RA Adachi J.; Aizawa K.; Akimura T.; Arikawa T.; Bono H.; Carninci P.;
 RA Fukuda S.; Furuno M.; Hanagaki T.; Hara A.; Hashizume W.;
 RA Hayashida K.; Hayatsu N.; Hiramoto K.; Hirooka T.; Hirozane T.;
 RA Hori F.; Imocani K.; Ishii Y.; Itoh M.; Kagawa I.; Kasukawa T.;
 RA Katoh H.; Kawai U.; Kojima Y.; Kondo S.; Kono H.; Kouda M.; Koya S.;
 RA Kurihara C.; Matsuyama T.; Miyazaki R.; Murata M.; Nakamura M.;
 RA Nishi K.; Nomura K.; Numazaki R.; Ohno M.; Ohsato N.; Okazaki Y.;
 RA Saito K.; Satoh H.; Sakai C.; Sakai K.; Sakazume N.; Sano H.;
 RA Sasaki D.; Shibata K.; Shinagawa A.; Shiraki T.; Sogabe Y.; Tagami M.;
 RA Tagawa A.; Takahashi F.; Takaku-Akahira S.; Takeda Y.; Tanaka T.;
 RA Tomaru A.; Toya T.; Yasunishi A.; Muramatsu M.; Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 DR EMBL; AK030375; BAC26930.1; -.
 DR MGD; MGI:1924104; Gpr30.
 DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0045028; F:purinergic nucleotide receptor activity, G- . . ; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. . . ; IEA.
 DR InterPro: IPR000276; GPCR Rhodopsn.
 DR InterPro: IPR002286; P2_purinoptor.
 DR Pfam: PF00001; 7cm.1.1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PRINTS: PR01157; P2YRNOCEPT.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SO SEQUENCE 375 AA; 42444 MW; A6E7D795C3A6AB05 CRC64;
 Query Match 12.9%; Score 223.5; DB 2; Length 375;
 Best Local Similarity 24.1%; Pred. No. 9.7e-09;
 Matches 83; Conservative 70; Mismatches 139; Indels 53; Gaps 16;
 QY 12 IVEELPACODQLGLSLGLVGVGVGLCYNALVLANLHRSKASMTMDVYFVNNAV 71
 DB 45 LIGDSEHQOYIAL-FLSCLYTIPLFPGFNGNLLIVNISPREKMTIPDLRYINLA 103
 QY 72 AGVLVSALAPVHLLGPSSRMALMSVGEVHVALQIPPNVSLVAMYSTAL---LSLDH 127
 DB 104 ADLIVADSLIEVFNLDQYDIAVLCTFMSLFIQI-----NMYSVPEFLTWMSFDR 155
 QY 128 YIERALPRYMASVYNTFH---VCGFWGAGALLTFSSILFYICSHVSTRALFCARQN 183
 DB 156 YL--ALAKMRGCLFRTKHARLSCGLIW--MASVSATLVFTVHLRHTEACFCRAD 210
 QY 184 AEAADATLVIGYVVP--ALATVALVLSRVREDTPLDRDTGLER---SAHRLVAT 238
 DB 211 VAEVQMELVTGLFIMPFAITGLCYSLIVRALIRAH-----RRRG-LRPRQALAMIRAV 264
 QY 239 VCTQGLMTPHYLILIGTVIISRGKPDV-----HYGLL-HFYVDFSLAFSSSF 290
 DB 265 VLVFFICWLPENVFISVH--LIQWTPQDPTPKQSFRAHYPLTGHIV---NLAISSNC 318
 QY 291 VNPILRYVNNQSPPSKLRIMKK---LPCGDHSCPDHMGVOQL 332
 DB 319 LNPILYFLEGTFRDLRLYVEQKTSLPALNRF--HATLKAVI 360
 RESULT 14
 OGHZNS PRELIMINARY; PRT; 355 AA.
 ID OGHZNS;
 AC OGHZNS;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Interleukin 8 receptor B CXCR2.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 ON NCBI_TaxID=9544;
 RN [1]
 RP SOURCE FROM N.A.
 RA Horlick R.A.; Zhao J.; Swanson R.N.; Webb M.L.; Strohl B.;
 RA Baldwin U.J.; Auld D.S.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 DR EMBL; AF540792; AN17318.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004918; F:interleukin-8 receptor activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0006935; P:chemotaxis; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . . ; IEA.
 DR InterPro: IPR000276; GPCR Rhodopsn.
 DR InterPro: IPR000057; IL8R_Rhodopsn.
 DR InterPro: IPR000174; IL8R_receptor.

DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPN.
 DR PRINTS; PR00427; INTRLEUKIN8R.
 DR PRINTS; PR00573; INTRLEUKIN8R.
 DR PROSITE; PS00237; G_PROTEIN_RECCEP_F1_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECCEP_F1_2; 1.
 KM G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 355 AA; 40292 MW; 6A9BF9A921F470C5 CRC64;

Query Match 12.0%; Score 208.5; DB 2; Length 355;
 Best Local Similarity 23.5%; Pred. No. 1.2e-07;
 Matches 84; Conservative 59; Mismatches 127; Indels 87; Gaps 13;

QY 1 MWSGSRNGTGVLELP-----ACQDLQGLS-----LLSLGLVGVGVG 41
 DB 9 LMGEDFSNYSYSDLPSPDPAVPCRPESLEINKYFVVIYALVFLSLG-----60
 QY 42 LCVNALLVNLHASKAMTPDVFVMAVAGVLSALAPVHLGPPSSRWALMSVGGEV 101
 DB 61 ---NSLVMLVILHSRVSITDYLNLAMADLFPALTLP-----WAAKVGWIF 108
 QY 102 H-----VALQIPFNVSILVAMYSTALLSDHYIERA-LPTYMASVYNTRHVCGFVWG 153
 DB 109 FGTFLCKVSLKENVFYSGILL--LACISVDRYLAIVHATRTLQKRYLVKFCISIGL 166
 QY 154 GALLTSFSSILFYICSHVSTRALECAKQNAEADATLVF-----IGYVPALATLYAL 207
 DB 167 LSLLLAPVLLFRRAVYPPYISPVCEYDMGNTAKMVMVRLIPQTGFIVPLIMLFCY 226
 QY 208 VLSRVREDTPLDRTGRLEPSAHRLLVATVCTQFGIMTPHYILGHVYIISR-----262
 DB 227 GLTLR-----TLFRAHMGQ-KHRAMRVIFAVVLIPLCLMPLHYLVADLTMRRLINET 280
 QY 263 -----GKPVDA-HYLGLLHFPVKDFSKLLAFSSSFVTPLLRYMNSFPSKQLQRLM 311
 DB 281 CORRNINIDQALDATEILGILH-----SCNPLIYAFIGQKFRHGLKLIL 324

RESULT 15
 Q8HZN3
 ID Q8HZN3 PRELIMINARY; PRT; 355 AA.
 AC Q8HZN3;
 DT 01-MAR-2003 (T:EMBLrel. 23, Created)
 DT 01-MAR-2003 (T:EMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (T:EMBLrel. 26, Last annotation update)
 DE Interleukin 8 receptor B CXCR2.
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 OC NCBI_TaxID=9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Horlick R.A., Zhao J., Swanson R.N., Webb M.L., Strohl B.,
 RA Baldwin J.J., And D.S.,
 RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
 CC - SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC - SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 DR EMBL; AF540794; AAI17320.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004818; F:interleukin-8 receptor activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR InterPro; IPR000057; IL8R_receptor.
 DR InterPro; IPR00174; IL8_receptor.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPN.
 DR PRINTS; PR00427; INTRLEUKIN8R.
 DR PRINTS; PR00573; INTRLEUKIN8R.
 DR PROSITE; PS00237; G_PROTEIN_RECCEP_F1_1; 1.

DR PROSITE; PS00262; G_PROTEIN_RECCEP_F1_2; 1.
 KM G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 355 AA; 40253 MW; 69DEF0C0D0717A53 CRC64;

Query Match 11.9%; Score 206.5; DB 2; Length 355;
 Best Local Similarity 23.5%; Pred. No. 1.8e-07;
 Matches 81; Conservative 61; Mismatches 138; Indels 65; Gaps 12;

QY 2 WSGSRNGTGVLELP-----ACQDLQGLS-----LLSLGLVGVGVG 42
 DB 10 WTGEDFSNYSYSDLPSPDPAVPCRPESLEINKYFVVIYALVFLSLG-----60
 QY 43 CVNALLVNLHASKAMTPDVFVMAVAGVLSALAPVHLGPPSSRWALMSVGGEV 102
 DB 61 ---NSLVMLVILHSRVSITDYLNLAMADLFPALTLP-----WAAKVGWIF 109
 QY 103 -----VALQIPFNVSILVAMYSTALLSDHYIERA-LPTYMASVYNTRHVCGFVWG 154
 DB 110 GFTFLCKVSLKENVFYSGILL--LACISVDRYLAIVHATRTLQKRYLVKFCISIGL 167
 QY 155 ALLTSFSSILFYICSHVSTRALECAKQNAEADATLVF-----IGYVPALATLYAL 208
 DB 168 LSLLLAPVLLFRRAVYPPYISPVCEYDMGNTAKMVMVRLIPQTGFIVPLIMLFCY 227
 QY 209 LLSRVREDTPLDRTGRLEPSAHRLLVATVCTQFGIMTPHYILGHVYIISR--GKPV 266
 DB 228 FTLR-----TLFRAHMGQ-KHRAMRVIFAVVLIPLCLMPLHYLVADLTMRRLINET 281
 QY 267 DAHYLGLLHFPVKDFSKLLAFSSSFVTPLLRYMNSFPSKQLQRLM 311
 DB 282 QRH--SDINQALDATEILGIFHSCNPLIYAFIGQKFRHGLKLIL 324

Search completed: December 14, 2004, 20:25:06
 Job time: 57.3673 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:34:50 ; Search time 90.94 Seconds

(without alignments)
14040.270 Million cell updates/sec

Title: US-09-995-225B-41

Perfect score: 27
Sequence: 1 gcgcctcgtgagcagcagcagcagccac 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb_ba: 2: gb_ptg: 3: gb_in: 4: gb_cm: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vi:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	27	6 AX498218	AX498218 Sequence
2	27	100.0	1130	6 AX375235	AX375235 Sequence
3	27	100.0	1167	9 AY635179	AY635179 Homo sapi
4	27	100.0	1526	6 AX664701	AX664701 Sequence
5	27	100.0	101882	9 AC021089	AC021089 Homo sapi
6	27	100.0	110000	2 AL139235_0	AL139235 Homo sapi
7	27	100.0	140885	2 AC119798	AC119798 Rattus no
8	27	100.0	183326	10 AC124438	AC124438 Mus muscu
9	27	100.0	202934	10 AC147502	AC147502 Mus muscu
10	27	100.0	230813	2 AP006494_2	AP006494 Rattus no
11	22.8	84.4	110000	6 BD183201	BD183201 Novel G p
12	22	81.5	32	6 BD103849	BD103849 Novel G p
13	22	81.5	1059	6 AX463227	AX463227 Sequence
14	21	77.8	1062	6 BD183203	BD183203 Novel G p
15	21	77.8	1062	6 BD183204	BD183204 Novel G p
16	21	77.8	1062	6 AX463235	AX463235 Sequence
17	21	77.8	1062	6 AX498192	AX498192 Sequence
18	21	77.8	1062	6 AX664703	AX664703 Sequence
19	21	77.8	1062	6 AX664703	AX664703 Sequence

20	21	77.8	1062	6 BD103851	BD103851 Novel G p
21	21	77.8	1062	6 BD103852	BD103852 Novel G p
22	21	77.8	1119	6 AX254975	AX254975 Sequence
23	21	77.8	1343	6 AX254977	AX254977 Sequence
24	20.6	76.3	846	11 BVL66828	BVL66828 TNFRSF14_1
25	20.2	74.8	1398	6 AX934641	AX934641 Sequence
26	20.2	74.8	181788	2 CR396589	CR396589 Dnlo rer
27	19.8	73.3	664	6 HSA341042	HSA341042 Homo sapi
28	19.8	73.3	3796	6 AX151499	AX151499 Sequence
29	19.8	73.3	10321	1 AE012251	AE012251 Xenomom
30	19.8	73.3	35815	1 LMF2802	LMF2802 Leishmani
31	19.8	73.3	130500	2 AC148130	AC148130 Carolina
32	19.8	73.3	170415	2 AC123299	AC123299 Rattus no
33	19.8	73.3	221641	3 LMF1046	LMF1046 Leishmani
34	19.8	73.3	292967	14 AP69029	AP69029 White spo
35	19.8	73.3	305107	6 AX151386	AX151386 Sequence
36	19.8	73.3	305107	14 AF440570	AF440570 White spo
37	19.8	73.3	307287	14 AF440570	AF440570 Shrimp wh
38	19.6	72.6	178	6 AX795059	AX795059 Sequence
39	19.6	72.6	8391	6 CO714847	CO714847 Sequence
40	19.6	72.6	8676	6 AX795161	AX795161 Sequence
41	19.6	72.6	8742	6 AX795155	AX795155 Sequence
42	19.6	72.6	10537	1 AE014779	AE014779 Blidobac
43	19.6	72.6	70384	9 AC100764	AC100764 Homo sapi
44	19.6	72.6	79677	8 AC090435	AC090435 Chlamydom
45	19.6	72.6	87340	1 AP005224	AP005224 Corynebac

ALIGNMENTS

RESULT 1	AX498218	Sequence 41 from Patent WO0242461.	27 bp	DNA	linear	PAT 26-SEP-2002
LOCUS	AX498218					
DEFINITION	AX498218					
ACCESSION	AX498218.1	GI:23343140				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
ORIGIN						
Query Match						
Best Local Similarity						
Matches						
Db						
Qy						
RESULT 2	AX375235	Sequence 25 from Patent WO0210387.	1130 bp	DNA	linear	PAT 01-MAR-2002
LOCUS	AX375235					
DEFINITION	AX375235					
ACCESSION	AX375235.1	GI:19169953				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R., Elliott, V.S., Ramkumar, J., Baughn, M.R., Kallio, D.A., Walla, N.K., Haffala, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L., Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.
G-protein coupled receptors
Patent: WO 0210387-A 25 07-FEB-2002;
Incyte Genomics, Inc. (US)
Location/Qualifiers

FEATURES
source
1. .1130
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474767CB1"

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 27; DB 6; Length 1130;
Pred. No. 0.5;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 GCGCTCATGAGCAGCAGCGCCAC 27
|||||
6 GCGCTCATGAGCAGCAGCGCCAC 32

RESULT 3
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY635179 1167 bp mRNA linear PRI 04-JUL-2004
Homo sapiens G protein-coupled receptor 139 (GPR139) mRNA, complete cds.
AY635179
AY635179.1 GI:49413551
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1167)
Bommer, T.I., Nagle, J.W. and Kauffman, D.
Complete coding sequence of GPR139
Unpublished
2 (bases 1 to 1167)
Bommer, T.I., Nagle, J.W. and Kauffman, D.
Direct Submission
Submitted (25-MAY-2004) Lab of Genetics, NIMH, Bldg 36, Rm 3D06, MSC4094, Bethesda, MD 20892-4094, USA
Location/Qualifiers

FEATURES
source
1. .1167
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p12.3"
/issue_type="brain"
1. .1167
/gene="GPR139"
92. .1153
/gene="GPR139"
/codon_start=1
/product="G protein-coupled receptor 139"
/protein_id="AA165818.1"
/db_xref="GI:49413552"
/translation="MEHTRAHLAANSSLSWSPGACGAGFVPPVYYSILLCGLPAN
ILTVILSQLVARROKSSYNLLAADAIDLPLFVFDLLEDILMMQPOVDPK
IIEVLFSSHTSIWITPLTIDRYIAVCHPKHTVSPARTKRVISVYITCFLTS
IPYVWPNIWTEIDYSTVHVLIVHCFVYQVCSIFFLINSIVYKLRKSFRL
RGYSTGKTALIFITISIFALIMAPRIIMILVHVGAPIONRWLHINSDIANMAL
NTAINFLYCFISKRFRMAATLKAFFKQKQPOVQFTNNHFSITSSPWISPSNSHC
IKMLVYQDKNGKPIKVPSP"
complement (1142..1167)

gene
CDS
primer_bind
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 1167;
Pred. No. 0.5;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 GCGCTCATGAGCAGCAGCGCCAC 27
|||||
86 GCGCTCATGAGCAGCAGCGCCAC 112

RESULT 4
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX664701 1526 bp DNA linear PAT 22-MAR-2003
Sequence 10 from Patent WO02074960.
AX664701
AX664701.1 GI:29164461
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Leiby, K.R., Kapeller-Liebermann, R. and Gluckmann, M.
38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
of human proteins and uses thereof
Patent: WO 02074960-A 10 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers

FEATURES
source
1. .1526
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
139..1200
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD80046.1"
/db_xref="GI:29164462"
/translation="MEHTRAHLAANSSLSWSPGACGAGFVPPVYYSILLCGLPAN
ILTVILSQLVARROKSSYNLLAADAIDLPLFVFDLLEDILMMQPOVDPK
IIEVLFSSHTSIWITPLTIDRYIAVCHPKHTVSPARTKRVISVYITCFLTS
IPYVWPNIWTEIDYSTVHVLIVHCFVYQVCSIFFLINSIVYKLRKSFRL
RGYSTGKTALIFITISIFALIMAPRIIMILVHVGAPIONRWLHINSDIANMAL
NTAINFLYCFISKRFRMAATLKAFFKQKQPOVQFTNNHFSITSSPWISPSNSHC
IKMLVYQDKNGKPIKVPSP"

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 27; DB 6; Length 1526;
Pred. No. 0.5;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
1 GCGCTCATGAGCAGCAGCGCCAC 27
|||||
133 GCGCTCATGAGCAGCAGCGCCAC 159

RESULT 5
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC021089 101862 bp DNA linear PRI 22-MAR-2003
Homo sapiens chromosome 16 clone CTD-2264D9, complete sequence.
AC021089
AC021089.8 GI:29150340
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 101862)
DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
Direct Submission
Unpublished
2 (bases 1 to 101862)
DOE Joint Genome Institute.

gene
CDS
primer_bind
ORIGIN

TITLE Direct Submission
JOURNAL Submitted (14-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 5 (bases 1 to 101882)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Mar 22, 2003 this sequence version replaced gi:20330797.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.sbgsc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated total Number of Errors is 0.
Location/Qualifiers
source
1. .101882
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2264D9"

ORIGIN
Query Match 100.0%; Score 27; DB 9; Length 101882;
Best Local Similarity 100.0%; Pred. No. 0.45; Mismatches 0; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCTCATGAGCAGCAGCAGCCAC 27
Db 60121 GCGCTCATGAGCAGCAGCAGCCAC 60095

RESULT 6
AL139235 0/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL139235 Accession AL139235
Fragment Name Begin End
AL139235_0 1 110000
AL139235_1 100001 210000
AL139235_2 200001 300001
AL139235_3 300001 407025
LOCUS AL139235 407025 bp DNA 1linear HTG 16-JUL-2001
DEFINITION Homo sapiens chromosome 9 clone RP11-3215, 31 unordered pieces.
ACCESSION AL139235
VERSION AL139235.22 GI:11034497
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE 1
AUTHORS Plumb, B.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
request: clonerequest@sanger.ac.uk
On Oct 26, 2000 this sequence version replaced gi:11024977.
----- Genome Center
Center: Sanger Centre
Center code: SC

Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: B31215
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 388894 bases at least Q40
Consensus quality: 395933 bases at least Q30
Consensus quality: 400050 bases at least Q20
Insert size: 404025; sum-of-contigs
Insert size: 174695; 17.7% error; agarose-fp
Quality coverage: 4.93x in Q20 bases; sum-of-contigs Quality
coverage: 14.49x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2425: contig of 2425 bp in length
* 2426 2425: gap of 100 bp
* 2526 10023: contig of 7498 bp in length
* 10024 10123: gap of 100 bp
* 10124 36364: contig of 26241 bp in length
* 36365 36464: gap of 100 bp
* 36465 43873: contig of 7409 bp in length
* 43874 43973: gap of 100 bp
* 43974 43988: contig of 6015 bp in length
* 43989 50088: gap of 100 bp
* 50089 93603: contig of 43515 bp in length
* 93604 93703: gap of 100 bp
* 93704 96191: contig of 2488 bp in length
* 96192 96291: gap of 100 bp
* 96292 96433: contig of 3142 bp in length
* 96434 99533: gap of 100 bp
* 99534 101556: contig of 2023 bp in length
* 101557 101656: gap of 100 bp
* 101657 184109: contig of 82453 bp in length
* 184110 184209: gap of 100 bp
* 184210 290696: contig of 106487 bp in length
* 290697 290796: gap of 100 bp
* 290797 292892: contig of 2096 bp in length
* 292893 292992: gap of 100 bp
* 292993 296018: contig of 3026 bp in length
* 296019 296118: gap of 100 bp
* 296119 322154: contig of 26036 bp in length
* 322155 322254: gap of 100 bp
* 322255 324746: contig of 2392 bp in length
* 324747 324746: gap of 100 bp
* 324747 329069: contig of 4323 bp in length
* 329070 329159: gap of 100 bp
* 329160 331371: contig of 2202 bp in length
* 331372 331471: gap of 100 bp
* 331472 333909: contig of 2438 bp in length
* 333910 334009: gap of 100 bp
* 334010 369445: contig of 35436 bp in length
* 369446 369545: gap of 100 bp
* 369546 371996: contig of 2451 bp in length
* 371997 372096: gap of 100 bp
* 372097 375925: contig of 3829 bp in length
* 375926 376025: gap of 100 bp
* 376026 378132: contig of 2107 bp in length
* 378133 378232: gap of 100 bp
* 378233 381590: contig of 3338 bp in length
* 381591 381690: gap of 100 bp
* 381691 383726: contig of 2036 bp in length
* 383727 383826: gap of 100 bp
* 383827 390832: contig of 7006 bp in length

FEATURES	Location/Qualifiers
*	390932: gap of 100 bp
*	390933
*	354216: contig of 3284 bp in length
*	394217
*	394316: gap of 100 bp
*	394317
*	396358: contig of 2042 bp in length
*	396359
*	396458: gap of 100 bp
*	398733
*	398739: contig of 2275 bp in length
*	398774
*	398833: gap of 100 bp
*	398834
*	402386: contig of 3552 bp in length
*	402386
*	402485: gap of 100 bp
*	404522: contig of 2037 bp in length
*	404523
*	404622: gap of 100 bp
*	404555
*	407025: contig of 2403 bp in length
*	404623

```

misc_feature
1. .2475
/note="assembly_fragment:05385
fragment_chain:1"
2526. .10023
/note="assembly_fragment:00331
fragment_chain:1"
10124. .36364
/note="assembly_fragment:06402
fragment_chain:1"
36465. .43873
/note="assembly_fragment:01436
fragment_chain:1"
43974. .4998
/note="assembly_fragment:04547
fragment_chain:1"
50089. .59603
/note="assembly_fragment:01636
fragment_chain:1"
93704. .96191
/note="assembly_fragment:02957
fragment_chain:2"
96292. .99433
/note="assembly_fragment:00499
fragment_chain:2"
99534. .10156
/note="assembly_fragment:03112
fragment_chain:2"
101657. .184109
/note="assembly_fragment:01419
fragment_chain:3"
184210. .290696
/note="assembly_fragment:00925
fragment_chain:3"
290797. .292892
/note="assembly_fragment:03065
fragment_chain:4"
292893. .296018
/note="assembly_fragment:05057
fragment_chain:4"
296119. .322154
/note="assembly_fragment:00269"
322255. .324646
/note="assembly_fragment:00686"
324747. .328069
/note="assembly_fragment:01113"
329170. .331371
/note="assembly_fragment:01461"
331472. .333909
/note="assembly_fragment:01516"
334010. .356945
/note="assembly_fragment:01942.0"
365946. .371196
/note="assembly_fragment:02194"

```

```
misc_feature 372097. 375925 /note="assembly_fragment:02249"
misc_feature 376026. 378132 /note="assembly_fragment:02485"
misc_feature 378233. 381590 /note="assembly_fragment:02649"
misc_feature 381691. 383726 /note="assembly_fragment:02869"
misc_feature 383827. 390832 /note="assembly_fragment:03540"
misc_feature 390933. 394216 /note="assembly_fragment:04405"
misc_feature 394317. 396358 /note="assembly_fragment:04595"
misc_feature 396459. 398733 /note="assembly_fragment:05768"
misc_feature 398834. 402385 /note="assembly_fragment:06106"
misc_feature 402486. 404522 /note="assembly_fragment:06204"
misc_feature 404623. 407025 /note="assembly_fragment:06292"

ORIGIN
```

Query Match	100.0%;	Score 27;	DB 2;	Length 110000;
Best Local Similarity	100.0%;	Freq. No. 0.45;		
Matches	27;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qx	1	GGCGCTCATGTAGACACACGCGCCAC	27	
Db	76457	GGCGCTCATGTAGACACACGCGCCAC	76431	

[illegible]

Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: <http://genome.imb-jena.de/>
Contact: gscj-submit@genome.imb-jena.de

Project Information
Center project name: RA23
Center clone name: RP13-339N19

Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329
 Consensus quality: 140504 bases at least Q40
 Consensus quality: 140589 bases at least Q30
 Consensus quality: 140641 bases at least Q20
 Quality coverage: 12.80x

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

- * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.
- * 1 27033: contig of 27033 bp in length
- * 27034 27133: gap of unknown length
- * 27134 67824: contig of 40691 bp in length
- * 67825 67924: gap of unknown length
- * 67925 140885: contig of 72961 bp in length.

FEATURES

source

```
1.140885
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/chromosome="1"
/map="q35"
/clone="RP32-329N19"
/clone_id="RP32-329N19"
/clone_id="RP32-329N19"
```

ORIGIN

Query Match 100.0%; Score 27; DB 2; Length 140885;
 Best Local Similarity 100.0%; Pred. No. 0.45;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGGAGCAGCAGCCGAC 27

Db 117161 GCGCTCATGGAGCAGCAGCCGAC 117135

RESULT 8
 AC124438 183326 bp DNA linear ROD 13-NOV-2003

DEFINITION Mus musculus BAC clone RP24-260K9 from chromosome 7, complete

AC124438

VERSION AC124438.4 GI:23462943

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 183326)

AUTHORS McPherson, J.D., and Waterston, R.H.

TITLE The sequence of Mus musculus BAC clone RP24-260K9

JOURNAL Unpublished (2001)

REFERENCE 2 (bases 1 to 183326)

AUTHORS Wilson, R.

TITLE Sequencing of Mus musculus

JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 183326)

AUTHORS McPherson, J.D., and Waterston, R.H.

TITLE Direct Submission

REFERENCE 4 (bases 1 to 183326)

AUTHORS McPherson, J.D., and Waterston, R.H.

TITLE Direct Submission

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park

Pathway, St. Louis, MO 63108, USA

4 (bases 1 to 183326)

McPherson, J.D., and Waterston, R.H.

Direct Submission

Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park

Pathway, St. Louis, MO 63108, USA

5 (bases 1 to 183326)

McPherson, J.D., and Waterston, R.H.

Direct Submission

Submitted (02-OCT-2002) Genome Sequencing Center, 4444 Forest Park

Pathway, St. Louis, MO 63108, USA

6 (bases 1 to 183326)

Wilson, R.

Direct Submission

Submitted (13-NOV-2003) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Oct 2, 2002 this sequence version replaced gi:22539265.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@wustl.wustl.edu

Center project name: M_BB0260K09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP32-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

Location/Qualifiers

```
1.183326
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/map="7"
/clone="RP24-260K9"
/clone_id="RP24-260K9"
/clone_id="RP24-260K9"
421..527
/rpt_family="L1"
/rpt_family="L1"
780..849
/rpt_family="L1"
/rpt_family="L1"
1229..1489
/rpt_family="L1"
/rpt_family="L1"
1925..2147
/rpt_family="MER1_type"
2269..2345
/rpt_family="B4"
```

```

repeat_region      2432..2639      /rpt_family="L1"
repeat_region      2657..2747      /rpt_family="L1"
repeat_region      2753..2819      /rpt_family="ID"
repeat_region      2825..2891      /rpt_family="L1"
repeat_region      3225..4557      /rpt_family="L1"
repeat_region      5367..5580      /rpt_family="L1"
repeat_region      6490..6682      /rpt_family="MIR"
repeat_region      6777..6881      /rpt_family="B2"
repeat_region      6889..7045      /rpt_family="Alu"
repeat_region      9156..9335      /rpt_family="B4"
repeat_region      10113..10156      /rpt_family="MER2_type"
repeat_region      10440..10701      /rpt_family="B4"
repeat_region      11049..11114      /rpt_family="ID"
repeat_region      11205..11581      /rpt_family="L1"
repeat_region      12639..12861      /rpt_family="B4"
repeat_region      12916..12988      /rpt_family="B4"
repeat_region      13314..13598      /rpt_family="B4"
repeat_region      14985..15253      /rpt_family="B4"
repeat_region      15284..15355      /rpt_family="MIR"
repeat_region      15669..15726      /rpt_family="ERV1"
repeat_region      16139..16229      /rpt_family="L1"
repeat_region      16418..16525      /rpt_family="L1"
repeat_region      16880..16932      /rpt_family="L1"
repeat_region      16942..17139      /rpt_family="MIR"
repeat_region      17302..17445      /rpt_family="L1"
repeat_region      17459..17508      /rpt_family="Alu"
repeat_region      17667..17787      /rpt_family="L1"
repeat_region      18127..18260      /rpt_family="Alu"
repeat_region      18386..18472      /rpt_family="L2"
repeat_region      18473..18818      /rpt_family="MIR"
repeat_region      18831..19229      /rpt_family="MALR"
repeat_region      19435..19530      /rpt_family="MALR"
repeat_region      20110..20262      /rpt_family="MIR"
repeat_region      20395..20447      /rpt_family="B4"
repeat_region      20458..20581      /rpt_family="ERV1"
repeat_region      21632..21854      /rpt_family="MALR"
repeat_region      21927..22288      /rpt_family="MER1_type"

```

```

repeat_region      /rpt_family="MALR"
repeat_region      23000..23102      /rpt_family="MIR"
repeat_region      23289..23480      /rpt_family="B2"
repeat_region      23481..23670      /rpt_family="B2"
repeat_region      23985..24034      /rpt_family="L2"
repeat_region      24461..24730      /rpt_family="L1"
repeat_region      25760..25970      /rpt_family="B4"
repeat_region      26406..26524      /rpt_family="MER1_type"
repeat_region      26748..27090      /rpt_family="Alu"
repeat_region      28283..28734      /rpt_family="MALR"
repeat_region      29104..29319      /rpt_family="ERV1"
repeat_region      29316..29425      /rpt_family="MER2_type"
repeat_region      29597..29747      /rpt_family="MER2_type"
repeat_region      29825..29939      /rpt_family="MIR"
repeat_region      30206..30399      /rpt_family="MIR"
repeat_region      31148..31285      /rpt_family="MER1_type"
repeat_region      32032..32336      /rpt_family="MER1_type"
repeat_region      /rpt_family="MALR"

```

```

Query Match      100.0%; Score 27; DB 10; Length 183326;
Best Local Similarity 100.0%; Pred. NO. 0.44; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;
QY      1 GCGCTCATGTGAGCACGACGACGCCAC 27
Db      66628 GCGCTCATGTGAGCACGACGACGCCAC 66654

```

```

RESULT 9
AC147502      202934 bp      DNA      linear      ROD 15-MAY-2004
LOCUS
DEFINITION Mus musculus BAC clone RP23-111N9 from chromosome 7, complete
sequence.
ACCESSION AC147502
VERSION AC147502.2 GI:46275996
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 202934)
Swearengen-Shahid,S., Shahid,S., Cotton,M., Kozlowski,A. and
Meyer,R.
TITLE The sequence of Mus musculus BAC clone RP23-111N9
UNPUBLISHED (2001)
2 (bases 1 to 202934)
AUTHORS Wilson,R.K.
DIRECT SUBMISSION
SUBMITTED (05-DEC-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 202934)
AUTHORS Wilson,R.K.
DIRECT SUBMISSION
SUBMITTED (08-APR-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 202934)
AUTHORS Wilson,R.K.
DIRECT SUBMISSION
TITLE

```


JOURNAL

Submitted (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
On Apr 8, 2004 this sequence version replaced gi:38708154.

COMMENT

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: M_BA0111N09

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RRC1-23 BAC Library has been constructed by Kazuhiro Osegawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC124438.

Location/Qualifiers

1. 202934

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="7"

/map="7"

/clone="RP23-111N9"

/clone_1b="RRC1-23"

1959. 2627

/rpt_family="L1"

2691. 3205

/rpt_family="L1"

3505. 3950

/rpt_family="L1"

4204. 4359

/rpt_family="L1"

5371. 5662

/rpt_family="B4"

5767. 6129

/rpt_family="BRV1"

6738. 6796

/rpt_family="BRV1"

6770. 6830

/rpt_family="BRV1"

6874. 6946

/rpt_family="BRV1"

6990. 7039

/rpt_family="BRV1"

7040. 7314

/rpt_family="B4"

7901. 7986

repeat_region
/rpt_family="BRV1"
8120. 8340
/rpt_family="MALR"

unsure

8363. 8545

/note="Sequence derived from one plasmid subclone."

11282. 11670

repeat_region

12408. 12698

repeat_region

13934. 14144

repeat_region

14154. 14294

repeat_region

15089. 15228

repeat_region

15259. 15531

repeat_region

15907. 16027

repeat_region

16070. 16180

repeat_region

17754. 17867

repeat_region

19213. 19396

repeat_region

19774. 20083

repeat_region

20358. 20449

repeat_region

20624. 20721

repeat_region

20791. 20941

repeat_region

21006. 21133

repeat_region

22495. 22732

repeat_region

22897. 22990

repeat_region

23022. 23100

repeat_region

23308. 23498

repeat_region

23850. 23934

repeat_region

25280. 25406

repeat_region

25682. 25852

repeat_region

26200. 26385

repeat_region

26659. 26898

repeat_region

28500. 28665

repeat_region

28668. 28810

repeat_region

29284. 29461

repeat_region

29469. 29542

repeat_region

31411. 31526

repeat_region

33813. 33868

repeat_region

33897. 34003

repeat_region

35786. 35962

repeat_region

35988. 36388

repeat_region

36433. 36586

repeat_region

```

repeat_region      36698..36835
                    /rpt_family="MRLR"
repeat_region      37663..37957
                    /rpt_family="L1"
repeat_region      37971..38090
                    /rpt_family="Alu"
repeat_region      38143..39094
                    /rpt_family="L1"
repeat_region      39871..40127
                    /rpt_family="RMR17C"
repeat_region      40375..40589
                    /rpt_family="B2"
repeat_region      40617..40715
                    /rpt_family="ID"
repeat_region      40978..41128
                    /rpt_family="MIR"
repeat_region      43035..43298
                    /rpt_family="L1"
repeat_region      43928..44131
                    /rpt_family="L2"
repeat_region      44422..44452
                    /rpt_family="L2"
repeat_region      44463..44862
                    /rpt_family="RMR17C"
repeat_region      44655..44914
                    /rpt_family="ERVX"

```

```

Query Match      100.0%; Score 27; DB 10; Length 202934;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 GCCTCATGAGACACGACGCCCCAC 27
Db      190319 GCCTCATGAGACACGACGCCCCAC 190345

```

```

RESULT 10
LOCUS      AC114097 230813 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-90E1, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC114097
VERSION    AC114097.5 GI:30580092
KEYWORDS   HTG; HTGS PHASE2; HTGS_DPAFT; HTGS_ENRICHED.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.

```

```

REFERENCE
AUTHORS    1 (bases 1 to 230813)
           Muzny,D,Marie, Metzker,M, Lee, Abramson, S., Adams, C., Alder, J.,
           Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angilano, D.,
           Anyadebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
           Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
           Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
           Bryant, N., Bubay, C., Burch, P., Burrell, K., Calderon, E.,
           Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
           Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, Y., Chen, Z., Chu, J.,
           Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
           Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Divya, K.,
           Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
           Egan, A., Baccotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
           Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
           Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
           Gebregeorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
           Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
           Harevey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
           Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
           Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, D., Jackson, A.,
           Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Joilivet, A.,
           Kapaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C.,
           Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
           Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

COMMENT

```

Lorenshewala, L., Loulseged, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, B.,
 Mlisaavijevic, A., Miner, G., Minja, E., Montenegro, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mundiada, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwankweme, O., Okunolu, G., Olarunbasogun, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Polinder, A., Popovic, D., Primus, B., Pu, L., L.,
 Puazo, M., Quintoz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sison, I., Stiller, C. D., Smajic, D.,
 Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wiczyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weiss, G., and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 230813)
 Worley, K. C.
 Direct Submission
 Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 230813)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:24819108.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GLIA
 Center clone name: CH230-90E1
 ----- Summary Statistics
 Assembly program: Atlas 3.0:
 Consensus quality: 226424 bases at least Q40
 Consensus quality: 227854 bases at least Q30
 Consensus quality: 228799 bases at least Q20
 Estimated insert size: 240103; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 230813: contig of 230813 bp in length.
Location/Qualifiers

FEATURES

source

1..230813
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-90B1"

ORIGIN

Query Match 100.0%; Score 27; DB 2; Length 230813;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGCTCATGGAGCAGCAGCCGAC 27
DB 90822 GCGCTCATGGAGCAGCAGCCGAC 90796

RESULT 11

AP006494.2/c

WPCOMMENT

Sequence split into 9 fragments LOCUS AP006494 Accession AP006494

Fragment Name	Begin	End
AP006494.0	1	110000
AP006494.1	100001	210000
AP006494.2	200001	310000
AP006494.3	300001	410000
AP006494.4	400001	510000
AP006494.5	500001	610000
AP006494.6	600001	710000
AP006494.7	700001	810000
AP006494.8	800001	849694

Continuation '3 of 9' of AP006494 from base 200001 (AP006494 Cyanidioschyzon merolae str.

Query Match 84.4%; Score 22.8; DB 2; Length 110000;
Best Local Similarity 92.3%; Pred. No. 29;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCGCTCATGGAGCAGCAGCCGAC 26
DB 64886 GCGCTCATGGAGCAGCAGCCGAC 64861

RESULT 12

BD183201

LOCUS BD183201 32 bp DNA linear PAT 17-JUN-2003
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD183201
VERSION BD183201.1 GI:31875401
KEYWORDS JP 2002345481-A/1.
SOURCE eSynthetic construct
ORGANISM eSynthetic construct
artificial sequences.
1 (bases 1 to 32)
Terao, Y., Matsui, H. and Shintani, Y.
Novel G protein-coupled receptor and its DNA
Patent: JP 2002345481-A 1 03-DEC-2002;
TAKEDA CHEMICAL INDUSTRIES LTD
OS Artificial Sequence
PN JP 2002345481-A/1
PD 03-DEC-2002
PR 01-JUN-2001 JP 2001166688
PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/09, A61K45/00, A61P1/00, A61P9/00, A61P25/00, PC
A61P29/00,
PC A61P35/00, A61P37/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, PC
C12N1/21,

COMMENT

AUTHORS Terao, Y., Matsui, H. and Shintani, Y.
TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: JP 2002345481-A 1 03-DEC-2002;
TAKEDA CHEMICAL INDUSTRIES LTD
OS Artificial Sequence
PN JP 2002345481-A/1
PD 03-DEC-2002
PR 01-JUN-2001 JP 2001166688
PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/09, A61K45/00, A61P1/00, A61P9/00, A61P25/00, PC
A61P29/00,
PC A61P35/00, A61P37/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, PC
C12N1/21,

PC C12N5/10, C12P21/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC
G01N33/53,
PC G01N33/566, C12N15/00, C12N5/00
CC Designed oligonucleotide primer to amplify DNA encoding TGR8
FH Key Location/Qualifiers
FT source 1..32
/organism="Artificial Sequence".
Location/Qualifiers

FEATURES

source

1..32
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 81.5%; Score 22; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 CATGAGCAGCAGCAGCCGAC 27
DB 6 CATGAGCAGCAGCAGCCGAC 27

RESULT 13

BD103849

LOCUS BD103849 32 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD103849
VERSION BD103849.1 GI:22649423
KEYWORDS WO 0194582-A/1.
SOURCE eSynthetic construct
ORGANISM eSynthetic construct
artificial sequences.
1 (bases 1 to 32)
Terao, Y., Matsui, H. and Shintani, Y.
Novel G protein-coupled receptor and its DNA
Patent: WO 0194582-A 1 13-DEC-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI
SHINTANI
OS Artificial Sequence
PN WO 0194582-A/1
PD 13-DEC-2001
PR 01-JUN-2001 WO 2001P004643
PR 02-JUN-2000 JP 00P 170446-23-JUN-2000 JP 00P 194926 PI
YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00,
PC A61P29/00,
PC A61P9/00, A61P35/00, A61P37/00, A61P1/00, G01N33/566, PC
G01N33/50,
PC G01N33/15
CC Designed oligonucleotide primer to amplify DNA encoding TGR8
FH Key Location/Qualifiers
FT source 1..32
/organism="Artificial Sequence".
Location/Qualifiers

1..32
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

COMMENT

Query Match 81.5%; Score 22; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 CATGAGCAGCAGCAGCCGAC 27
DB 6 CATGAGCAGCAGCAGCCGAC 27

RESULT 14
AX463227
LOCUS AX463227 1059 bp DNA linear PAT 15-JUL-2002

Db 1 ATGAGCAGCAGCGCCAC 21
Search completed: December 15, 2004, 00:42:54
Job time : 92.94 secs

DEFINITION Sequence 1 from Patent WO0248358.
ACCESSION AX463227
VERSION AX463227.1 GI:21886199
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Smolyar,A., Zhu,Z., Encinas,J., Watanabe,S. and Okigami,H.
TITLE Regulation of human chemokine-like receptor
JOURNAL Patent: WO 0248358-A 1 20-JUN-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..1059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 77.8%; Score 21; DB 6; Length 1059;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCGCCAC 27
Db 1 ATGAGCAGCAGCGCCAC 21

RESULT 15
BD183203 1062 bp DNA linear PAT 17-JUN-2003
LOCUS BD183203
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD183203
VERSION BD183203.1 GI:31875403
KEYWORDS JP 2002345481-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Terao,Y., Matsui,H. and Shintani,Y.
TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: JP 2002345481-A 3 03-DEC-2002;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT OS Homo sapiens (human)
FN JP 2002345481-A/3
PD 03-DEC-2002
PF 01-JUN-2001 JP 2001166688
PI YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/09,A61K45/00,A61P1/00,A61P3/00,A61P9/00,A61P25/00, PC
A61P29/00,
PC A61P35/00,A61P37/00,C07K14/705,C07K16/28,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC
G01N33/53,
PC G01N33/566,C12N15/00,C12N5/00
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
/organism="Homo sapiens (human)".
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 77.8%; Score 21; DB 6; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCGCCAC 27

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:00:04 ; Search time 12.6451 Seconds
(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225B-41

Perfect score: 27
Sequence: 1 ggcgcacatggagacacgcacgcacccac 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001a:*\n5: geneseqn2001b:*\n6: geneseqn2002a:*\n7: geneseqn2002b:*\n8: geneseqn2003a:*\n9: geneseqn2003b:*\n10: geneseqn2003c:*\n11: geneseqn2003d:*\n12: geneseqn2004e:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	ABT04896	ABT04896 Human G-P
2	27	100.0	1130	AAD29672	AAD29672 Human G-P
3	27	100.0	1202	ABV73364	ABV73364 Human TGR
4	27	100.0	1526	AAD46858	AAD46858 Human 7TM
5	27	100.0	1526	ACA60889	ACA60889 Human CDN
6	27	100.0	1526	ABSS57024	ABSS57024 cDNA enco
7	27	100.0	2189	ADD18120	ADD18120 Human G-P
8	27	100.0	2198	ACFO5275	ACFO5275 Human G-P
9	27	100.0	2272	AD028955	AD028955 Human nov
10	22	81.5	32	ABK15560	ABK15560 Novel G-P
11	22	81.5	42	ADJ87864	ADJ87864 G-coupled
12	22	81.5	42	AD056045	AD056045 Human NOV
13	21	77.8	25	AD030468	AD030468 Human nov
14	21	77.8	28	ACC71787	ACC71787 Human nov
15	21	77.8	116	ACC71795	ACC71795 Human G-P
16	21	77.8	1038	ABV73373	ABV73373 Mouse TGR
17	21	77.8	1038	AD028957	AD028957 Mouse nov
18	21	77.8	1059	ABN84269	ABN84269 Human che
19	21	77.8	1062	ABK15562	ABK15562 cDNA enco
20	21	77.8	1062	ABK15563	ABK15563 cDNA enco
21	21	77.8	1062	ABT04873	ABT04873 Human G-P

22	21	77.8	1062	6	ABN84273	ABN84273 Human che
23	21	77.8	1062	10	ACC44115	ACC44115 Human AXO
24	21	77.8	1062	10	ACC71785	ACC71785 Human G-P
25	21	77.8	1110	9	ACF05276	ACF05276 Human G-P
26	21	77.8	1119	4	AA515730	AA515730 DNA enco
27	21	77.8	1119	10	ADJ87766	ADJ87766 G-coupled
28	21	77.8	1119	12	ADJ79322	ADJ79322 NOV10a co
29	21	77.8	1119	12	AD056001	AD056001 DNA enco
30	21	77.8	1343	4	AA515731	AA515731 DNA enco
31	21	77.8	1343	4	ADJ87768	ADJ87768 G-coupled
32	21	77.8	1343	12	ADJ79324	ADJ79324 NOV11 cod
33	21	77.8	1343	12	AD056003	AD056003 DNA enco
34	21	77.8	1776	10	ADF70587	ADF70587 Ophan re
35	20.2	74.8	1398	10	ADG33813	ADG33813 Actinomy
36	19.8	73.3	3796	4	AAH62741	AAH62741 Shrimp wh
37	19.8	73.3	305107	4	AAH62689	AAH62689 Shrimp wh
38	19.6	72.6	178	10	ADJ81026	ADJ81026 INSP010 m
39	19.6	72.6	8676	10	ADJ81128	ADJ81128 INSP010 m
40	19.6	72.6	8742	10	ADJ81122	ADJ81122 INSP010 n
41	19.6	72.6	8838	6	ABV93367	ABV93367 Human NOV
42	19.6	72.6	8862	10	ADJ81136	ADJ81136 INSP010 P
43	19.6	72.6	34980	6	ABO81849	ABO81849 Bifidobac
44	19.6	72.6	34980	6	ABO81848	ABO81848 Bifidobac
45	19	70.4	612	4	AAH67672	AAH67672 Human HTL

ALIGNMENTS

RESULT 1
ABT04896
ID ABT04896 standard; DNA; 27 BP.
XX AC ABT04896;
XX DT 11-OCT-2002 (first entry)
XX DE Human G protein coupled receptor hrup35 PCR primer SEQ ID NO: 41.
XX KW Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31;
XX KM hrup32; hrup33; hrup34; hrup35; hrup36; hrup37; PCR; primer; ss.
XX OS Homo sapiens.
XX XX
XX PN WO200242461-A2.
XX PD 30-MAY-2002.
XX PF 26-NOV-2001; 2001WO-US044386.
XX PR 27-NOV-2000; 2000US-0253404P.
XX PR 12-DEC-2000; 2000US-0255366P.
XX PR 20-FEB-2001; 2001US-0270286P.
XX PR 20-FEB-2001; 2001US-0270286P.
XX PR 06-APR-2001; 2001US-0282032P.
XX PR 06-APR-2001; 2001US-0282356P.
XX PR 06-APR-2001; 2001US-0282358P.
XX PR 06-APR-2001; 2001US-0282365P.
XX PR 14-MAY-2001; 2001US-0290917P.
XX PR 31-JUL-2001; 2001US-0309208P.
XX PA (AREN-) ARENA PHARM INC.
XX PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX WPI; 2002-565565/60.
XX DR Novel endogenous and non-endogenous versions of G protein-coupled
XX PT receptor useful for identification of candidate compounds as receptor
XX PT agonists or antagonists for use as therapeutic agents.
XX PS Example 1; Page 28; 84pp; English.

CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a PCR primer used to isolate a GPCR coding sequence of the invention
XX
SQ Sequence 27 BP; 6 A; 12 C; 7 G; 2 T; 0 U; 0 Other;
Query Match 100.0%; Score 27; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCGCTCATGAGCAGCAGCGCCAC 27
Db 1 GCGCTCATGAGCAGCAGCGCCAC 27
RESULT 2
AAD29672 standard; cDNA; 1130 BP.
XX
AC AAD29672;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human G-protein coupled receptor (GCRC-6) cDNA.
XX
KW Human; G-protein coupled receptor; GCRC-6; cell proliferative disorder;
KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;
KW metabolic; hepatic; psoriasis; cancer; epilepsy; Alzheimer's disease;
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cyclostatic;
KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
KW nootropic; neuroprotective; cardiant; immunosuppressive; anorectic;
KW virocidic; gene; ss.
XX
XX Homo sapiens.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 12..1130
FT /*tag= a
FT /*product= "Human GCRC-6"
FT sig_peptide 12..152
FT /*tag= b
FT mat_peptide 153..1127
FT /*tag= c
FT /*product= "Human mature GCRC-6"
XX
XX MO200210387-A2.
XX
XX 07-FEB-2002.
XX
XX 25-JUL-2001; 2001MO-US023433.
XX
XX 27-JUL-2000; 2000US-0221478P.
XX 03-AUG-2000; 2000US-022368P.
XX 21-AUG-2000; 2000US-0227054P.
XX 08-SEP-2000; 2000US-0231121P.
XX 13-SEP-2000; 2000US-0232243P.
XX 15-SEP-2000; 2000US-0232591P.
XX 22-SEP-2000; 2000US-0235146P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Thornton M, Patterson C, Lai P, Burford N, Yue H, Gandhi AR;
PI Elliot VS, Ramkumar J, Baughn MR, Kallik DA, Wala NK, Hafalia ADA;
PI Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L, Graul RC;
XX Warren BA, Lee EA, Ding L;
XX
XX MPI; 2002-188744/24.
XX
XX P-PSDB; AAE18645.
DR

XX
XX New human G-protein coupled receptor polypeptide for diagnosis,
PT prevention and treatment of cell proliferative, neurological,
PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic
PT disorders.
XX
XX Claim 5; Page 141; 150pp; English.
XX
XX The invention relates to novel human G-protein coupled receptors (GCRC)
CC and their encoding polynucleotides. GCRC is useful as an immunogen for
CC preparing monoclonal and polyclonal antibodies. GCRC is useful for
CC diagnosing, treating and preventing a cell proliferative disorder (e.g.,
CC hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,
CC Alzheimer's disease, Huntington's disease, Parkinson's
CC disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,
CC myocardial infarction), gastrointestinal disorder (e.g., gastritis,
CC cirrhosis, Crohn's disease), an autoimmune/inflammatory disorder (e.g.,
CC acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,
CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,
CC osteoporosis), and viral infections. GCRC is useful in a number of drug
CC screening techniques, and to analyse the proteome of a tissue or cell
CC type. GCRC is useful for creating knockin humanised animals or
CC transgenic animals to model human diseases, in somatic or germline gene
CC therapy, to generate a transcript image of a tissue or cell type, for
CC detecting differences in the chromosomal location due to translocation,
CC inversion, etc., among normal, carrier or affected individuals, and as
CC hybridization probes for mapping naturally occurring genomic sequences.
CC GCRC is useful in Southern or Northern analysis, dot blot or other
CC membrane-based technologies, in PCR technologies, in dipstick, pin,
CC microformat enzyme linked immunosorbant (ELISA)-like assays, and in
CC microarrays utilising fluids or tissues from patients to detect altered
CC GCRC expression. The present sequence is human GCRC-6 cDNA
XX
SQ Sequence 1130 BP; 258 A; 357 C; 226 G; 289 T; 0 U; 0 Other;
Query Match 100.0%; Score 27; DB 6; Length 1130;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCGCTCATGAGCAGCAGCGCCAC 27
Db 6 GCGCTCATGAGCAGCAGCGCCAC 32
RESULT 3
ID ABV73364 standard; DNA; 1202 BP.
XX
XX ABV73364;
XX
XX 22-JAN-2003 (first entry)
XX
XX Human TGR20 polypeptide encoding DNA.
XX
XX G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;
XX antileptic; nootropic; neuroprotective; antianemic; antileuc; human;
XX antiparkinsonian; antileptic; TGR20; gene; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 68..1129
FT /*tag= a
FT /*product= "TGR20"
FT /*note= "GPCR polypeptide"
XX
XX MO200277001-A2.
XX
XX 03-OCT-2002.
XX
XX 08-MAR-2002; 2002MO-US007171.
XX
XX 09-MAR-2001; 2001US-00802803.
XX
XX

PR 16-MAR-2001; 2001US-0276649P.
XX
XX (TULUA-) TULARIK INC.
XX
XX Tian H, Zhao J, Chen J, Cutler G;
XX
XX WPI; 2003-018881/01.
DR P-PSDB; ABB82499.
XX
XX
PT New G-protein coupled receptor polypeptides and polynucleotides useful
PT for identifying compounds for treating a TGR-associated disorder, e.g.
PT psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
PT disease, anemia.
XX
XX
PS Claim 6; Page 63; 87pp; English.
XX
XX The invention relates to G-protein coupled receptor (GPCR) polypeptides
CC and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
CC TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
CC useful for identifying compounds for treating a TGR-associated disorder,
CC such as psoriasis, inflammatory bowel disease, hyperlipidemia,
CC Parkinson's disease, Huntington's disease, anemia, immune and blood
CC disorders, ulcerative colitis, Crohn's disease or spleen enlargement.
CC They are also useful for identifying cells such as kidney, liver,
CC hypothalamus, colon, adipose, or spleen cells, for forensics and
CC paternity determination, diagnosing diseases and examining signal
CC transduction. The present sequence represents a human TGR20 polypeptide
CC encoding DNA
XX
SQ Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 27; DB 8; Length 1202;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GCGCTCATGAGACACGACGCCAC 27
Db 62 GCGCTCATGAGACACGACGCCAC 88
XX
RESULT 4
AAD46858
ID AAD46858 standard; cDNA; 1526 BP.
XX
XX AAD46858;
AC
XX
DT 27-JAN-2003 (first entry)
XX
DB Human 7TM domain receptor 65507 cDNA.
XX
XX Human; adenosine deaminase; seven transmembrane domain receptor; cancer;
KW 7TM; glycoprotease; immune disorder; IGA deficiency; allergy; arrhythmia;
KW rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;
KW hypertension; ischemic heart disease; obesity; myocardial infarction;
KW endometrial cell disorder; Grave's disease; psoriasis; brain disorder;
KW Parkinson's disease; Alzheimer's disease; haematopoietic disorder;
KW cerebral oedema; metabolic disorder; liver disorder; platelet disorder;
KW chromosome mapping; tissue typing; gene therapy; neuroprotective;
KW cytoskeletal; anorectic; cardiant; haemostatic; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 139..1200
FT CDS
FT /*tag= a
FT /product= "Human 65507 protein"
FT /note= "This region is specifically claimed as SEQ ID NO:
FT 12 in claim 1 of the specification"
FT 139..306
FT /*tag= b
FT 307..1197
FT /*tag= c
FT /product= "Human mature 65507 protein"

XX
XX PN WO200274960-A2.
XX
XX PD 26-SEP-2002.
XX
XX 08-NOV-2001; 2001WO-US051427.
XX
XX PF 08-NOV-2000; 2000US-0246768P.
XX
XX PR 08-NOV-2000; 2000US-0246772P.
XX
XX PR 15-NOV-2000; 2000US-0249185P.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
XX PI Leiby KR, Kapeller-Libermann R, Glucksmann M;
XX
XX DR WPI; 2002-759898/82.
XX
XX DR P-PSDB; AAE29236.
XX
XX PT New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,
PT useful for diagnosing and treating cancer, immune, cardiovascular,
PT hematopoietic, brain, pain, metabolic, liver or platelet disorders, and
PT in pharmacogenomics.
XX
XX
XX PS Claim 1; Fig 17; 178pp; English.
XX
XX CC The present invention relates to novel 38650, 28472, 5495, 65507, 81588
CC or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-
CC protease or seven transmembrane domain (7TM) receptor family members.
CC Sequences of the invention are useful in diagnosing and treating cancer
CC or aberrant cellular proliferation and/or differentiation (e.g. colon or
CC lung cancer), immune disorders (e.g. selective IGA deficiency, rheumatoid
CC arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g.,
CC hypertension, atherosclerosis, arrhythmias, ischemic heart disease,
CC myocardial infarction, thrombus) including endothelial cell disorders
CC (e.g. psoriasis, Grave's disease), haematopoietic disorders, brain
CC pain and metabolic disorders (e.g. obesity), liver disorders or platelet
CC disorders. They are also useful in screening assays, predictive medicine
CC (e.g. diagnostic assays, prognostic assays, monitoring clinical trials
CC and pharmacogenetics) and prophylactic and therapeutic methods. The
CC nucleic acids may also be used in chromosome mapping, tissue typing and
CC forensic biology and as surrogate markers. Sequences of the invention are
CC also used in gene therapy. The present sequence is human 7TM domain
CC receptor 65507 cDNA
XX
XX SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
XX
Query Match 100.0%; Score 27; DB 6; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GCGCTCATGAGACACGACGCCAC 27
Db 133 GCGCTCATGAGACACGACGCCAC 159
XX
RESULT 5
ACA60889
ID ACA60889 standard; cDNA; 1526 BP.
XX
XX ACA60889;
AC
XX
DT 08-JUL-2003 (first entry)
XX
DB Human cDNA encoding 65507 protein.
XX
XX Human; ss; gene; cancer; aberrant cellular proliferation;
KW differentiation; immune disorders; heart disorder; brain disorder;
KW cardiovascular disorder; endometrial cell disorder; pain disorder;
KW haematopoietic disorder; blood vessel disorder; metabolic disorder;
KW liver disorder; platelet disorder.
XX
XX Homo sapiens.
OS

```
XX Key Location/Qualifiers
FH CDS 139..1200
FT /tag= a
FT /product= "Protein 65507"
FT /note= "This CDS is specifically claimed in claim 1"
XX
XX US2003009017-A1.
XX
XX 09-JUN-2003.
XX
XX 08-NOV-2001; 2001US-00012140.
XX
XX 08-NOV-2000; 2000US-0246768P.
XX 08-NOV-2000; 2000US-0246772P.
XX 15-NOV-2000; 2000US-0249185P.
XX
XX (LEIB/) LEIBY K R.
XX (KAPE/) KAPPELLER-LIBERMANN R.
XX (GLUC/) GLUCKSMANN M A.
XX
XX Leiby KR, Kapeller-libermann R, Glucksmann MA;
XX
XX WPI; 2003-428888/40.
XX P-PSDB; ABU09571.
XX
XX New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
XX molecules, useful for diagnosing, treating cancer, pain, or immune,
XX heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
XX and liver disorders.
XX
XX Claim 2; Fig 17; 90pp; English.
XX
XX The invention relates to an isolated 38650 (encoding adenosine
XX deaminase), 28472 (encoding a glycoprotease), 5495 (encoding a 7
XX transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
XX a sequence which is at least 60% identical to the six nucleic acids or
XX their open reading frames, fragments of at least 15 nucleotides,
XX naturally occurring variants, or a DNA insert of the plasmid deposited
XX with the American Type Culture Collection as Accession No. not defined in
XX the specification, which encodes the amino acid sequence). Also included
XX are a host cell containing the nucleic acids (used to produce the
XX proteins), the encoded proteins, an antibody that selectively binds to
XX the polypeptide, and identifying a compound that binds to/modulates the
XX activity of the polypeptide. The nucleic acid molecules, polypeptides and
XX methods are useful for: diagnosing, treating cancer, aberrant cellular
XX proliferation and/or differentiation, immune disorders, heart disorders,
XX cardiovascular disorders including endothelial cell disorders,
XX haematopoietic disorders, blood vessel disorders, brain disorders, pain
XX and metabolic disorders, liver disorders and platelet disorders (many
XX examples of these disorders are given in the specification). The present
XX sequence is the Human CDNA 65507
XX
XX Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 27; DB 8; Length 1526;
XX Best Local Similarity 100.0%; Pred. No. 0.11;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
XX
XX Cancer; aberrant cell proliferation; aberrant cell differentiation;
XX breast cancer; ovarian cancer; prostate cancer; colon cancer;
XX lung cancer; immune disorder; heart disorder; cardiovascular disorder;
XX endothelial disorder; hematopoietic disorder; blood vessel disorder;
XX brain disorder; pain; metabolic disorder; liver disorder; diabetes;
XX platelet disorder; carcinoma; sarcoma; leukemia; Hodgkin's disease;
XX autoimmune disorder; hypertension; atherosclerosis; heart failure;
XX myocardial infarction; ischemic heart disease; Crohn's disease;
XX Grave's disease; Kawasaki syndrome; Raynaud's disease; aneurysm;
XX cerebral ischemia; peripheral neuropathy; Alzheimer's disease;
XX Parkinson's disease; anorexia nervosa; cachexia; gene; ss;
XX 7 transmembrane receptor.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 139..1200
FT /tag= a
FT /product= "7 transmembrane receptor 65507"
FT /note= "Specifically claimed in claim 1"
XX
XX WO200277233-A2.
XX
XX 03-OCT-2002.
XX
XX 08-NOV-2001; 2001WO-US046724.
XX
XX 08-NOV-2000; 2000US-0246768P.
XX 08-NOV-2000; 2000US-0246772P.
XX 15-NOV-2000; 2000US-0249185P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Leiby KR, Kapeller-libermann R, Glucksmann M;
XX
XX WPI; 2003-029938/02.
XX P-PSDB; ABG71163.
XX
XX New adenosine deaminase, glycoprotease and seven transmembrane domain
XX nucleic acids and polypeptides, designated 38650, 28472, 5495, 65507,
XX 81588 and 14354, useful for treating e.g. leukemia, Hodgkin's disease or
XX hypertension.
XX
XX Claim 2; Fig 17; 178pp; English.
XX
XX The invention describes isolated 38650, 28472, 5495, 65507, 81588 and
XX 14354 nucleic acid molecules (I) and their encoded polypeptides (II). The
XX 38650 nucleic acid molecule comprises a sequence encoding adenosine
XX deaminase. The 28472 nucleic acid molecule comprises a sequence encoding
XX a human glycoprotease. The 5495, 65507, 81588 and 14354 comprise
XX sequences that encode a human seven transmembrane domain (7TM). The
XX 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid and polypeptide
XX sequences are useful for diagnosing, preventing or treating a subject
XX with or at risk of developing a disorder, e.g. cancer or aberrant
XX cellular proliferation and/or differentiation (e.g. breast, ovarian,
XX prostate, colon or lung cancer), immune disorders, heart disorders,
XX cardiovascular disorders, endothelial disorders, hematopoietic disorders,
XX blood vessel disorders, brain disorders, pain and metabolic disorders,
XX liver disorders or platelet disorders. These disorders include carcinoma,
XX sarcoma, leukemia, Hodgkin's disease, autoimmune disorders,
XX hypertension, atherosclerosis, heart failure, myocardial infarction,
XX ischemic heart disease, Crohn's disease, Grave's disease, Kawasaki
XX syndrome, Raynaud's disease, aneurysm, cerebral ischemia, peripheral
XX neuropathy, Alzheimer's disease, Parkinson's disease, anorexia nervosa,
XX cachexia or diabetes. This sequence encodes the novel human 7
XX transmembrane receptor 65507
XX
XX Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 27; DB 10; Length 1526;
XX Best Local Similarity 100.0%; Pred. No. 0.11;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Qy 1 GCGCTCATGAGCAGCAGCCGCAC 27
 ||||||||||||||||||||||||||||
 Db 133 GCGCTCATGAGCAGCAGCCGCAC 159

RESULT 7

ADD18120
 ID ADD18120 standard; DNA; 2189 BP.

AC ADD18120;
 XX
 DT 15-JAN-2004 (first entry)

DE Human G-protein coupled receptor (GPCR) DNA sequence Seq ID6.

XX G protein coupled receptor; GPCR; signal transduction pathway; G protein;
 KM Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;
 KM colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;
 KM sleeplessness; hypertension; anxiety; stress; renal failure;
 KM cardiovascular disorder; neural disorder; oncology disorder;
 KM immune disorder; neuroprotective; gene therapy; gene; da.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 1047..2157

FT /*tag= a
 /product= "Human G-protein coupled receptor"
 /transl_except= (2085..2091, aa:KPIKS)

PN W02003016478-A2.

XX 27-FEB-2003.

PD 15-AUG-2002; 2002MO-US026017.

PF 20-AUG-2001; 2001US-0313658P.

PR 12-SEP-2001; 2001US-0318675P.

PR 30-OCT-2001; 2001US-0340703P.

PR 26-NOV-2001; 2001US-0333417P.

PR 06-DEC-2001; 2001US-0338367P.

PR 06-FEB-2002; 2002US-0355596P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

PA Feder JN, Ramanathan CS, Gopal S, Mintier GA;

XX WPI; 2003-278558/27.

DR P-PSDB; ADD18023.

XX New nucleic acid, useful for manufacturing a medicament for preventing,

PT treating or ameliorating a medical condition e.g., neural disorder.

XX Example 2; SEQ ID NO 6; 251pp; English.

PS This invention relates to novel G protein coupled receptors (GPCRs) and
 XX their encoding nucleotide sequences. Many medically significant
 CC biological processes are mediated by proteins participating in signal
 CC transduction pathways involving G proteins. GPCRs are one of the largest
 CC receptor superfamilies known. These receptors are biologically important
 CC and malfunction of these receptors results in diseases such as
 CC Alzheimer's, Parkinson's, diabetes, dwarfism, colour blindness, retinal
 CC pigmentosa and asthma. They are also involved in depression,
 CC schizophrenia, sleeplessness, hypertension, anxiety, stress, renal
 CC failure and other cardiovascular, neural, oncology and immune disorders.
 CC A modulator of the GPCRs of the invention may have neuroprotective
 CC activity whilst the sequences of the invention may be useful for gene
 CC therapy. The invention may also be useful for manufacturing a medicament
 CC for preventing, treating or ameliorating a medical condition. The present
 CC sequence is the DNA sequence encoding a human GPCR of the invention.
 XX Sequence 2189 BP; 451 A; 691 C; 581 G; 466 T; 0 U; 0 Other;

Query Match 100.0%; Score 27; DB 10; Length 2189;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGCTCATGAGCAGCAGCCGCAC 27
 ||||||||||||||||||||||||||||
 Db 1041 GCGCTCATGAGCAGCAGCCGCAC 1067

RESULT 8

ACF05275
 ID ACF05275 standard; cDNA; 2198 BP.

AC ACF05275;
 XX
 DT 06-NOV-2003 (first entry)

DE Human G-protein coupled receptor HGRPMY34 cDNA.

XX HGRPMY34; G-protein coupled receptor; receptor; GPCR-P14; GPCR-145;
 KM human; neuroprotective; nootropic; tranquilizer; antiallergic;
 KM neuroleptic; antianemic; antidepressant; anticonvulsant; antiparkinsonian;
 KM cytoskeletal; cardiac; hypotensive; antianginal; analgesic; anorectic;
 KM anti-HIV; antisthmatic; osteopathic; uropathic; antitumor; antiallergic;
 KM gene therapy; gene; ss.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 1247..2165

FT /*tag= a
 /product= "HGRPMY34"

PN W02003050256-A2.

XX 19-JUN-2003.

PD 06-DEC-2002; 2002MO-US039290.

PF 06-DEC-2001; 2001US-0338371P.

PR (BRIM) BRISTOL-MYERS SQUIBB CO.

PA Feder JN, Gopal S, Mintier GA, Ramanathan CS;

XX WPI; 2003-577295/54.

DR P-PSDB; ABR62521.

XX New nucleic acid molecule encoding a human G-protein coupled receptor,

PT HGRPMY34, useful for diagnosing, preventing or treating diseases

PT involving the receptor, for example Parkinson's disease, dementia,

PT asthma, hypertension or cancer.

XX Claim 1; Fig 1A-B; 112pp; English.

PS The present sequence is that of cDNA encoding human HGRPMY34, a newly
 XX identified G-protein coupled receptor (GPCR) belonging to the group of
 CC 'Class A' GPCRs and showing homology to the Pfam model 7 transmembrane
 CC receptor, rhodopsin family. The cDNA was isolated by screening expressed
 CC sequence tag databases using GPCR sequences. HGRPMY34, also referred to
 CC as GPCR-P14 and/or GPCR-145, is highly expressed in brain (amygdala,
 CC caudate nucleus, corpus callosum, hippocampus, thalamus, substantia
 CC nigra), spinal cord and pituitary, indicating an association in
 CC neurological systems and conditions. It is also expressed in the bone
 CC marrow and testis. The invention provides HGRPMY34 polynucleotides,
 CC polypeptides and antibodies, expression vectors, host cells and antisense
 CC molecules, methods for screening for modulators of HGRPMY34 activity
 CC and/or function, and methods for diagnosing, treating, preventing and
 CC screening for disorders and diseases associated with abnormal HGRPMY34
 CC activity, including a disorder related to aberrant G-protein coupled
 CC signalling; a disorder related to aberrant cell cycle regulation;
 CC neurological disorders; anxiety; headache; migraine; schizophrenia; manic

depression; delirium; dementia; severe mental retardation and
 CC dyskinesias, such as Huntington's disease or Gilles de la Tourette's
 CC syndrome; Parkinson's disease; brain disorders; spinal cord disorders;
 CC affective disorders; neoplastic disorders; cardiovascular disorders;
 CC acute heart failure; hypotension; hypertension; angina pectoris;
 CC myocardial infarction; an immunological disorder; immune-related
 CC disorders; endocrine diseases; growth disorders; neuropathic pain;
 CC obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis
 CC ; psychosis; metabolic disorders; pituitary disorders; urinary retention;
 CC ulcers; allergies; or benign prostatic hypertrophy (all claimed)
 XX

Sequence 2198 BP, 456 A; 693 C; 581 G; 468 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 27; DB 9; Length 2198;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGGAGCAGCAGCCGCCAC 27
 Db 1041 GCGCTCATGGAGCAGCAGCCGCCAC 1067

RESULT 9
 ADO28955
 ID ADO28955 standard; cDNA; 2273 BP.
 AC ADO28955;
 XX
 DT 29-JUN-2004 (first entry)
 XX
 DE Human novel GPCR PCR3 polynucleotide, SEQ ID NO:54.
 XX

G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KM transgenic mouse; neurological disorder; adrenal gland disorder;
 KM colon disorder; intestinal disorder; cardiovascular disorder;
 KM muscular disorder; blood disorder; immune disorder;
 KM kidney disorder; metabolic disorder; nutritive disorder; cancer;
 KM joint disorder; uterus disorder; lung disorder; breast disorder;
 KM ovary disorder; prostate disorder; pancreas disorder; spleen disorder;
 KM skin disorder; stomach disorder; antiparkinsonian; antianemic;
 KM thymus disorder; thyroid disorder; antidiarrhoeic; antiarrhythmic;
 KM cytoskeletal; antiinflammatory; vasotropic; antidiabetic;
 KM CNS; central nervous system; vasorelaxant; antidiabetic;
 KM verucide; hepatotropic; antibacterial; antianemic; antidiabetic;
 KM dermatological; antitumor; antihypertensive; antidiabetic;
 KM immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
 KM gene; ss.
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO2004040000-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 XX
 PR 09-APR-2003; 2003US-0461329P.
 XX
 PA (PRIM-) PRIMAL INC.
 XX
 PI Galanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
 DR WPI; 2004-390329/36.
 DR P-Psdb; ADO28954.
 XX
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 PS Claim 13; SEQ ID NO 54; 542pp; English.

XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acid encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridise to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardioid infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
 CC nucleic acid of the invention. Note: The full sequence data for this
 CC patent did not form part of the printed specification; those sequences
 CC not shown were obtained in electronic format directly from WIP0 at
 CC ftp.wipo.int/pub/published_ptc_sequences.
 XX

Sequence 2273 BP, 503 A; 691 C; 549 G; 530 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 27; DB 12; Length 2273;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGGAGCAGCAGCCGCCAC 27
 Db 447 GCGCTCATGGAGCAGCAGCCGCCAC 473

RESULT 10
 ABR15560
 ID ABR15560 standard; DNA; 32 BP.
 XX
 AC ABR15560;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Novel G-protein coupled receptor TGR8, PCR primer TGR8sa1.
 XX

G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
 KM cytoskeletal; antiinflammatory; antitumor; fetal brain;
 KM central nervous system disease; circulatory organ disorder; cancer;
 KM metabolic disease; immunological disease; gastrointestinal disease;
 KM gene therapy; transgenic animal; human; PCR; primer; ss; TGR8sa1.
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO200194582-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 01-JUN-2001; 2001WO-JP004643.
 XX
 PR 02-JUN-2000; 2000JP-00170446.
 XX
 PR 23-JUN-2000; 2000JP-00194926.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX

PI Terao Y, Matsui H, Shintani Y;
 XX
 DR WPI; 2002-164317/21.
 XX
 PT Human fetal brain-originated G protein-coupled receptor protein TGR8 and
 PT encoding DNA, for developing drugs to treat e.g. diseases of the central
 PT nervous system or circulatory organs, cancer, and metabolic diseases.
 XX
 PS Example 2; Page 82; 102pp; Japanese.
 XX
 CC The invention describes a human fetal brain-originated G protein-coupled
 CC receptor protein, or its salt. The protein and encoded DNA are useful for
 CC developing drugs to treat e.g. diseases of the central nervous system or
 CC circulatory organs, cancer, metabolic diseases, immunological diseases
 CC and gastrointestinal diseases. The invention also describes creation of a
 CC probe for gene therapy and construction of a transgenic animal. This
 CC sequence represents the PCR primer TGR8sal used in the isolation of the
 CC novel human G-protein coupled receptor, TGR8
 XX
 SQ Sequence 32 BP; 7 A; 14 C; 8 G; 3 T; 0 U; 0 Other;
 Query Match 81.5%; Score 22; DB 6; Length 32;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 6 CATGAGCACACCGACGCCAC 27
 Db 6 CATGAGCACACCGACGCCAC 27
 RESULT 11
 ADJ87864
 ID ADJ87864 standard; DNA; 42 BP.
 AC ADJ87864;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE G-coupled protein receptor-related protein PCR primer #75.
 XX
 KM novel protein; G-coupled protein receptor-related protein;
 KM cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
 KM metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
 KM Huntington's disease; epilepsy; anxiety; pain; hypercholesterolaemia;
 KM obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
 KM viral infections; bacterial infection; parasitic infection;
 KM hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
 KM Alzheimer's disease; tuberculous sclerosis; hypercalcaemia; cerebral palsy;
 KM PCR; ss; primer.
 XX
 OS unidentified.
 XX
 PN WO2002102321-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 18-JUN-2002; 2002MO-US019522.
 XX
 PR 18-JUN-2001; 2001US-0298994P.
 PR 18-JUN-2001; 2001US-0299134P.
 PR 04-OCT-2001; 2001US-00972446.
 PR 06-JUN-2002; 2002US-00299134.
 PR 07-JUN-2002; 2002US-00298994.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
 PI Pena CE, Spaderna SK, Zhong M;
 XX WPI; 2003-167441/16.
 DR
 XX New MOLX polypeptides and polymucleotides, useful in gene therapy,
 PT particularly for treating or preventing e.g. cardiomyopathy,

PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
 PT cancer, stroke or pain.
 XX
 PS Example 10; SEQ ID NO 298; 378pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of novel G-
 CC coupled protein receptor-related (MOL) proteins. The DNA and protein
 CC sequences of the invention are useful for treating or preventing a MOL-
 CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
 CC associated with cell signal processing and metabolic pathway modulation,
 CC or diabetes. The DNA and protein sequences are also useful for the
 CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
 CC pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,
 CC systemic lupus erythematosus, viral infections, bacterial infections,
 CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or
 CC cerebral palsy. The present DNA sequence represents a PCR primer that was
 CC used in the exemplification of the invention.
 XX
 SQ Sequence 42 BP; 9 A; 20 C; 10 G; 3 T; 0 U; 0 Other;
 Query Match 81.5%; Score 22; DB 10; Length 42;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 6 CATGAGCACACCGACGCCAC 27
 Db 12 CATGAGCACACCGACGCCAC 33
 RESULT 12
 ADO56045
 ID ADO56045 standard; DNA; 42 BP.
 AC ADO56045;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Human NOV 11, PCR primer #17.
 XX
 KM human; cancer; obesity; diabetic nephropathy; acute pancreatitis; stroke;
 KM multiple sclerosis; ss; primer; PCR.
 XX
 OS Homo sapiens.
 XX
 PN US2004058862-A1.
 XX
 PD 25-MAR-2004.
 XX
 PF 18-SEP-2002; 2002US-00246583.
 XX
 PR 18-SEP-2002; 2002US-00246583.
 XX
 PA (MAJU/) MAJUMDER K.
 XX
 PI Majumder K;
 XX
 DR WPI; 2004-268835/25.
 XX
 PT Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic
 PT nephropathy, acute pancreatitis, strokes and multiple sclerosis.
 XX
 PS Example 9; Page 83; 87pp; English.
 XX
 CC The invention relates to novel isolated NOVX nucleic acids and encoded
 CC polypeptides. The nucleic acids, polypeptides and antibodies raised
 CC against the polypeptides are useful for preventing or treating diseases
 CC associated with aberrant NOVX expression or activity e.g., cancer,
 CC obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple
 CC sclerosis. The present sequence represents a PCR primer used to clone
 CC NOVX nucleic acids of the invention.
 XX
 SQ Sequence 42 BP; 9 A; 20 C; 10 G; 3 T; 0 U; 0 Other;

Query Match 81.5%; Score 22; DB 12; Length 42;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CATGAGCAGCAGCGCCAC 27
DB 12 CATGAGCAGCAGCGCCAC 33

RESULT 13
AD030468
AD030468 standard; DNA; 25 BP.

AC AD030468;
XX
XX
XX 29-JUL-2004 (first entry)
XX
XX
XX Human novel GPCR PGR3 RT-PCR primer, SEQ ID NO:1571.

KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW joint disorder; metabolic disorder; immune disorder; bone disorder;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; cancer;
KW cytostatic; antiinflammatory; vasotropic; antiparkinsonian; antiangiogenic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW vitucide; hepatotropic; antibacterial; antianemic; antiseborrhoeic;
KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW PGR3; reverse transcription-PCR; RT-PCR; primer; ss.
XX
XX Homo sapiens.
XX
XX MO200404000-A2.
XX
XX 13-MAY-2004.
XX
XX 09-SEP-2003; 2003MO-US028226.
XX
XX 09-SEP-2002; 2002US-0409303P.
XX 09-APR-2003; 2003US-0461329P.
XX
XX (PRIM-) PRIMAL INC.
XX
XX Galenaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,
XX Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H,
XX WPI; 2004-390329/36.
XX
XX Novel mammalian G protein coupled receptors, useful for identifying
XX PT compounds that modulates diagnosing and treating disease condition
XX PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
XX PT pectoris, Parkinson's disease.
XX
XX PS Disclosure; SEQ ID NO 1571; 542pp; English.

CC The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors

CC comprising a GPCR nucleic acid. The GPCR nucleic acid and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders (e.g., autoimmune disorders or
CC anaemia or leukaemia); immune disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a PCR primer
CC used in the isolation of cDNA encoding the novel human GPCR PGR3. Note:
CC The full sequence data for this patent did not form part of the printed
CC specification; those sequences not shown were obtained in electronic
CC format directly from WIP0 at ftp.wipo.int/pub/published_pcr_sequences.
XX
XX Sequence 25 BP; 6 A; 11 C; 6 G; 2 T; 0 U; 0 Other;

Query Match 77.8%; Score 21; DB 12; Length 25;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCAGCGCCAC 27
DB 1 ATGAGCAGCAGCAGCGCCAC 21

RESULT 14
ACCT1787
ID ACCT1787 standard; DNA; 28 BP.
XX
XX ACCT1787;
XX
XX 25-JUL-2003 (first entry)
XX
XX Human G protein-coupled receptor related DNA fragment #SEQ ID 5.
XX
XX Human; G protein-coupled receptor; anorectic; eating disorder; obesity;
XX ds.
XX
XX Homo sapiens.
XX
XX OS
XX MO2003027142-A1.
XX
XX 03-APR-2003.
XX
XX 19-SEP-2002; 2002MO-JP009626.
XX
XX 21-SEP-2001; 2001JP-00288278.
XX
XX (YAMA) YAMANOUCHI PHARM CO LTD.
XX
XX Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T,
XX Hiayama H,
XX WPI; 2003-333291/31.
XX
XX New G protein-coupled receptor and encoded gene, useful in screening
XX PT preventives or remedies for eating disorders or obesity.
XX
XX Example 1; Page 39; 48pp; Japanese.

CC The invention relates to a novel G protein-coupled receptor. The protein
CC and its encoded gene are useful for screening preventives or remedies for
CC eating disorders or obesity. The current sequence represents a G protein
CC coupled receptor related DNA fragment
XX
XX Sequence 28 BP; 7 A; 12 C; 7 G; 2 T; 0 U; 0 Other;

Query Match 77.8%; Score 21; DB 10; Length 28;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 ATGAGCAGCAGCAGCCGAC 27
 |||||
 1 ATGAGCAGCAGCAGCCGAC 21
 |||||

RESULT 15

ACC71795
 ID ACC71795 standard; DNA; 116 BP.
 XX
 AC ACC71795;
 XX
 DT 25-JUL-2003 (first entry)
 XX
 DE Human G protein-coupled receptor related PCR primer #SEQ ID 13.
 XX
 KW Human; G protein-coupled receptor; anorectic; eating disorder; obesity;
 KM PCR; primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003027142-A1.
 XX
 PD 03-APR-2003.
 XX
 PF 19-SEP-2002; 2002WO-JP009626.
 XX
 PR 21-SEP-2001; 2001JP-00288278.
 XX
 PA (YAMA) YAMANOUCHI PHARM CO LTD.
 XX
 PI Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T;
 PI Hiayama H;
 XX
 DR WPI; 2003-333291/31.
 XX
 PT New G protein-coupled receptor and encoded gene, useful in screening
 PT preventives or remedies for eating disorders or obesity.
 XX
 PS Example 5; Page 41; 48pp; Japanese.
 XX
 CC The invention relates to a novel G protein-coupled receptor. The protein
 CC and its encoded gene are useful for screening preventives or remedies for
 CC eating disorders or obesity. The current sequence represents a G protein
 CC coupled receptor related PCR primer sequence
 XX
 SQ Sequence 116 BP; 32 A; 34 C; 27 G; 23 T; 0 U; 0 Other;

Query Match 77.8%; Score 21; DB 10; Length 116;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 ATGAGCAGCAGCAGCCGAC 27
 |||||
 89 ATGAGCAGCAGCAGCCGAC 109
 |||||

Search completed: December 14, 2004, 22:43:55
 Job time : 13.6451 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 14, 2004, 22:27:19 ; Search time 2.27127 Seconds
(Without alignments)
8449.603 Million cell updates/sec

Title: US-09-995-225B-41

Perfect score: 27
Sequence: 1 ggcgcacatgagcagcagcagccac 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB. seq: *
2: /cgn2_6/prodata/1/ina/5B COMB. seq: *
3: /cgn2_6/prodata/1/ina/5A COMB. seq: *
4: /cgn2_6/prodata/1/ina/6B COMB. seq: *
5: /cgn2_6/prodata/1/ina/6B COMB. seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	70.4	720	3 US-08-913-014A-4	Sequence 4, Appl 1
C 2	19	70.4	720	4 US-09-653-285-4	Sequence 4, Appl 1
C 3	19	70.4	810	3 US-09-072-893C-5	Sequence 5, Appl 1
C 4	19	70.4	1017	4 US-09-027-287-38	Sequence 38, Appl 1
C 5	19	70.4	1017	4 US-09-252-656B-38	Sequence 38, Appl 1
C 6	19	70.4	1017	4 US-09-523-323-38	Sequence 38, Appl 1
C 7	19	70.4	1159	4 US-09-027-287-1	Sequence 1, Appl 1
C 8	19	70.4	1159	4 US-09-252-656B-1	Sequence 1, Appl 1
C 9	19	70.4	1159	4 US-09-523-323-1	Sequence 1, Appl 1
C 10	19	70.4	1353	3 US-08-913-014A-6	Sequence 6, Appl 1
C 11	19	70.4	1353	4 US-09-653-285-6	Sequence 6, Appl 1
C 12	19	70.4	1491	3 US-08-913-014A-5	Sequence 5, Appl 1
C 13	19	70.4	1491	4 US-09-653-285-5	Sequence 5, Appl 1
C 14	18.6	68.9	1533	4 US-09-270-767-12691	Sequence 12691, A
C 15	18.2	67.4	1458	4 US-10-011-200-1	Sequence 2, Appl 1
C 16	18.2	67.4	2346	4 US-10-011-200-2	Sequence 2, Appl 1
C 17	18	66.7	438	3 US-09-060-756-600	Sequence 600, App
C 18	18	66.7	438	4 US-09-670-314-600	Sequence 600, App
C 19	18	66.7	492	3 US-09-060-756-587	Sequence 587, App
C 20	18	66.7	492	4 US-09-670-314-587	Sequence 587, App
C 21	18	66.7	1423	1 US-08-386-729A-6	Sequence 3824, Ap
C 22	18	66.7	2259	4 US-09-489-039A-3024	Sequence 2, Appl 1
C 23	18	66.7	4403765	3 US-09-103-840A-2	Sequence 2, Appl 1
C 24	18	66.7	4403765	3 US-09-103-840A-2	Sequence 2, Appl 1
C 25	18	66.7	4411529	3 US-09-103-840A-1	Sequence 1, Appl 1
C 26	18	66.7	4411529	3 US-09-103-840A-1	Sequence 1, Appl 1
C 27	17.6	65.2	909	4 US-08-913-016C-1	Sequence 1, Appl 1

C 28	17.6	65.2	3588	1 US-08-197-792-32	Sequence 32, Appl 1
C 29	17.6	65.2	3588	1 US-08-459-850-32	Sequence 32, Appl 1
C 30	17.6	65.2	3588	1 US-08-459-214-32	Sequence 32, Appl 1
C 31	17.4	64.4	223	3 US-08-205-697A-14	Sequence 14, Appl 1
C 32	17.4	64.4	223	3 US-08-702-525-14	Sequence 14, Appl 1
C 33	17.4	64.4	223	4 US-09-837-867A-14	Sequence 14, Appl 1
C 34	17.4	64.4	223	5 PCT-US95-02576-14	Sequence 14, Appl 1
C 35	17.4	64.4	393	4 US-09-252-991A-9425	Sequence 9425, Ap
C 36	17.4	64.4	453	4 US-09-252-991A-10826	Sequence 10826, A
C 37	17.4	64.4	891	4 US-09-252-991A-9486	Sequence 9486, Ap
C 38	17.4	64.4	957	4 US-09-252-991A-9486	Sequence 9486, Ap
C 39	17.4	64.4	1081	3 US-09-372-422A-33	Sequence 7011, Ap
C 40	17.4	64.4	1125	4 US-09-252-991A-7011	Sequence 7011, Ap
C 41	17.4	64.4	1188	4 US-09-252-991A-5131	Sequence 5131, Ap
C 42	17.4	64.4	1261	3 US-08-205-697A-12	Sequence 12, Appl 1
C 43	17.4	64.4	1261	3 US-08-702-525-12	Sequence 12, Appl 1
C 44	17.4	64.4	1261	4 US-09-837-867A-12	Sequence 12, Appl 1
C 45	17.4	64.4	1261	5 PCT-US95-02576-12	Sequence 12, Appl 1

ALIGNMENTS

RESULT 1
US-08-913-014A-4/C
; Sequence 4, Application US/08913014A
; Patent No. 6235878
; GENERAL INFORMATION:
; APPLICANT: Nishii, Kazumori
; APPLICANT: Hikiuchi, Yukiko
; APPLICANT: Shintani, Yasushi
; TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS
; TITLE OF INVENTION: PRODUCTION AND USE
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: David G. Conlin, Esq.
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston,
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,014A
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP97/02480
; FILING DATE: July 17, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: David G. Conlin
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 342/47694
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: cDNA
; US-08-913-014A-4
Query Match 70.4%; Score 19; DB 3; Length 720;
Best Local Similarity 81.5%; Pred. No. 24;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGGAGCAGCAGCGCCGAC 27
Db 669 GCGTTCATCCAGCAGCAGCGACGAC 643

RESULT 2

US-09-653-285-4/C
; Sequence 4, Application US/09653285
; Patent No. 6590090

GENERAL INFORMATION:
; APPLICANT: Nishii, Kazumori

; Shintani, Yasushi
; TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS

PRODUCTION AND USE

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESSES:
; ADDRESSEE: David G. Conlin, Esq.

; DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP
; STREET: 130 Water Street

; CITY: Boston,
; STATE: MA

; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/653, 285

; FILING DATE: 31-Aug-2000
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/913,014

; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

; NAME: David G. Conlin
; REGISTRATION NUMBER: 27,026

; REFERENCE/DOCKET NUMBER: 342/47694
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440

; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
; LENGTH: 720 base pairs

TYPE: Nucleic acid
; STRANDEDNESS: Double

TOPOLOGY: Linear
; MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-653-285-4

Query Match 70.4%; Score 19; DB 4; Length 720;
Best Local Similarity 81.5%; Pred. No. 24;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGGAGCAGCAGCGCCGAC 27
Db 669 GCGTTCATCCAGCAGCAGCGACGAC 643

RESULT 3

US-09-072-993C-5/C
; Sequence 5, Application US/09072993C
; Patent No. 6346388

GENERAL INFORMATION:
; APPLICANT: Michael R. Brigham-Burke

; APPLICANT: Peter R. Young
; TITLE OF INVENTION: A METHOD OF IDENTIFYING AGONIST AND

; TITLE OF INVENTION: ANTAGONISTS FOR TUMOR NECROSIS RELATED RECEPTORS TR1 AND TR2
; FILE REFERENCE: GH-50030

CURRENT APPLICATION NUMBER: US/09/072,993C
; CURRENT FILING DATE: 1998-05-06

PRIOR APPLICATION NUMBER: 60/055,513
; PRIOR FILING DATE: 1997-08-13

PRIOR APPLICATION NUMBER: 60/056,980
; PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/057,550
; PRIOR FILING DATE: 1997-08-29

NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 5
; LENGTH: 810

TYPE: DNA
; ORGANISM: HOMO SAPIENS

US-09-072-993C-5

Query Match 70.4%; Score 19; DB 3; Length 810;
Best Local Similarity 81.5%; Pred. No. 24;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGGAGCAGCAGCGCCGAC 27
Db 743 GCGTTCATCCAGCAGCAGCGACGAC 717

RESULT 4

US-09-027-287-38/C
; Sequence 38, Application US/09027287A
; Patent No. 6479254

GENERAL INFORMATION:
; APPLICANT: Eder, Reinhard

; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ruben, Steven M.

; APPLICANT: Ullrich, Stephen
; TITLE OF INVENTION: Apoptosis Inducing Molecule II

; FILE REFERENCE: 1488.0650004
; CURRENT APPLICATION NUMBER: US/09/027,287A

; FILING DATE: 1998-02-20
; CURRENT FILING DATE: 1998-01-07

EARLIER APPLICATION NUMBER: US 08/822,953
; EARLIER FILING DATE: 1997-03-21

EARLIER APPLICATION NUMBER: US 60/030,157
; EARLIER FILING DATE: 1996-10-31

EARLIER APPLICATION NUMBER: US 60/013,923
; EARLIER FILING DATE: 1996-03-22

NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 38
; LENGTH: 1017

TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: CDS

LOCATION: (1)..(624)
US-09-027-287-38

Query Match 70.4%; Score 19; DB 4; Length 1017;
Best Local Similarity 81.5%; Pred. No. 25;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGGAGCAGCAGCGCCGAC 27
Db 573 GCGTTCATCCAGCAGCAGCGACGAC 547

RESULT 5

US-09-252-656B-38/C
; Sequence 38, Application US/09252656B
; Patent No. 6495520

GENERAL INFORMATION:
; APPLICANT: Peter R. Young

APPLICANT: Ebner, Reinhard
APPLICANT: Yu, Guo-Liang
APPLICANT: Ruben, Steven M.
APPLICANT: Zhang, Jun
APPLICANT: Ullrich, Stephen
APPLICANT: Zhai, Yifan
TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
FILE REFERENCE: 1488, 0650006
CURRENT APPLICATION NUMBER: US/09/252,656B
CURRENT FILING DATE: 1999-02-19
PRIOR APPLICATION NUMBER: US 60/075,409
PRIOR FILING DATE: 1998-02-20
PRIOR APPLICATION NUMBER: US 09/027,287
PRIOR FILING DATE: 1998-02-20
PRIOR APPLICATION NUMBER: US 09/003,886
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: US 08/822,953
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: US 60/013,923
PRIOR FILING DATE: 1996-03-22
PRIOR APPLICATION NUMBER: US 60/030,157
PRIOR FILING DATE: 1996-10-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 38
LENGTH: 1017
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(624)
OTHER INFORMATION: CDNA
US-09-252-656B-38

Query Match 70.4%; Score 19; DB 4; Length 1017;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTATGAGCAGCAGCGCCAC 27
DB 573 GCGTTCATCCAGCAGCAGCGCCAC 547

RESULT 6
US-09-523-323-38/c
Sequence 38, Application US/09523323
Patent No. 6635743
GENERAL INFORMATION:
APPLICANT: Ebner, Reinhard
APPLICANT: Yu, Guo-Liang
APPLICANT: Ruben, Steven M.
APPLICANT: Ullrich, Stephen
APPLICANT: Zhai, Yifan
TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
FILE REFERENCE: 1488, 065000C
CURRENT APPLICATION NUMBER: US/09/523,323
CURRENT FILING DATE: 2000-03-10
EARLIER APPLICATION NUMBER: 60/168,380
EARLIER FILING DATE: 1999-12-02
EARLIER APPLICATION NUMBER: 60/148,326
EARLIER FILING DATE: 1999-08-11
EARLIER APPLICATION NUMBER: 60/142,657
EARLIER FILING DATE: 1999-07-06
EARLIER APPLICATION NUMBER: 60/137,457
EARLIER FILING DATE: 1999-06-04
EARLIER APPLICATION NUMBER: 60/124,041
EARLIER FILING DATE: 1999-03-11
EARLIER APPLICATION NUMBER: 09/252,656
EARLIER FILING DATE: 1999-02-19
EARLIER APPLICATION NUMBER: 60/075,409
EARLIER FILING DATE: 1998-02-20
EARLIER APPLICATION NUMBER: 09/027,287
EARLIER FILING DATE: 1998-02-20

EARLIER APPLICATION NUMBER: 09/003,886
EARLIER FILING DATE: 1998-01-07
EARLIER APPLICATION NUMBER: 08/822,953
EARLIER FILING DATE: 1997-03-21
EARLIER APPLICATION NUMBER: 60/013,923
EARLIER FILING DATE: 1996-03-22
EARLIER APPLICATION NUMBER: 60/030,157
EARLIER FILING DATE: 1996-10-31
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 38
LENGTH: 1017
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(624)
US-09-523-323-38

Query Match 70.4%; Score 19; DB 4; Length 1017;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTATGAGCAGCAGCGCCAC 27
DB 573 GCGTTCATCCAGCAGCAGCGCCAC 547

RESULT 7
US-09-027-287-1/c
Sequence 1, Application US/09027287A
Patent No. 6479254
GENERAL INFORMATION:
APPLICANT: Ebner, Reinhard
APPLICANT: Yu, Guo-Liang
APPLICANT: Ruben, Steven M.
APPLICANT: Ullrich, Stephen
TITLE OF INVENTION: Apoptosis Inducing Molecule II
FILE REFERENCE: 1488, 0650004
CURRENT APPLICATION NUMBER: US/09/027,287A
CURRENT FILING DATE: 1998-02-20
EARLIER APPLICATION NUMBER: US 09/003,886
EARLIER FILING DATE: 1998-01-07
EARLIER APPLICATION NUMBER: US 08/822,953
EARLIER FILING DATE: 1997-03-21
EARLIER APPLICATION NUMBER: US 60/030,157
EARLIER FILING DATE: 1996-10-31
EARLIER APPLICATION NUMBER: US 60/013,923
EARLIER FILING DATE: 1996-03-22
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1169
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (49)..(768)
US-09-027-287-1

Query Match 70.4%; Score 19; DB 4; Length 1169;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTATGAGCAGCAGCGCCAC 27
DB 717 GCGTTCATCCAGCAGCAGCGCCAC 691

RESULT 8
US-09-252-656B-1/c
Sequence 1, Application US/09252656B
Patent No. 6495520

```
/ GENERAL INFORMATION:
/ APPLICANT: Ebner, Reinhard
/ APPLICANT: Yu, Guo-liang
/ APPLICANT: Ruben, Steven M.
/ APPLICANT: Zhang, Jun
/ APPLICANT: Ulrich, Stephen
/ APPLICANT: Zhai, Yifan
/ TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
/ FILE REFERENCE: 1488.0650006
/ CURRENT APPLICATION NUMBER: US/09/252,656B
/ PRIORITY FILING DATE: 1999-02-19
/ PRIOR APPLICATION NUMBER: US 60/075,409
/ PRIOR FILING DATE: 1998-02-20
/ PRIOR APPLICATION NUMBER: US 09/027,287
/ PRIOR FILING DATE: 1998-02-20
/ PRIOR APPLICATION NUMBER: US 09/003,886
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: US 08/822,953
/ PRIOR FILING DATE: 1997-03-21
/ PRIOR APPLICATION NUMBER: US 60/013,923
/ PRIOR FILING DATE: 1996-03-22
/ PRIOR APPLICATION NUMBER: US 60/030,157
/ PRIOR FILING DATE: 1996-10-31
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1169
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (49)..(768)
/ OTHER INFORMATION: DNA (cDNA)
US-09-252-656B-1

Query Match      70.4% Score 19; DB 4; Length 1169;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 GCGCTCATGAGCAGCAGCAGCCAC 27
DB      717 GCGTTCATCCAGCAGCAGCAGCCAC 691

RESULT 9
US-09-523-323-1/c
/ Sequence 1, Application US/09523323
/ Patent No. 6635743
/ GENERAL INFORMATION:
/ APPLICANT: Ebner, Reinhard
/ APPLICANT: Yu, Guo-liang
/ APPLICANT: Ruben, Steven M.
/ APPLICANT: Ulrich, Stephen
/ APPLICANT: Zhai, Yifan
/ TITLE OF INVENTION: Apoptosis Inducing Molecule II and Methods of Use
/ FILE REFERENCE: 1488.065000C
/ CURRENT APPLICATION NUMBER: US/09/523,323
/ PRIORITY FILING DATE: 2000-03-10
/ PRIOR APPLICATION NUMBER: 60/168,380
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: 60/148,326
/ PRIOR FILING DATE: 1999-08-11
/ PRIOR APPLICATION NUMBER: 60/142,657
/ PRIOR FILING DATE: 1999-07-06
/ PRIOR APPLICATION NUMBER: 60/137,457
/ PRIOR FILING DATE: 1999-06-04
/ PRIOR APPLICATION NUMBER: 60/124,041
/ PRIOR FILING DATE: 1999-03-11
/ PRIOR APPLICATION NUMBER: 09/252,656
/ PRIOR FILING DATE: 1999-02-19
/ PRIOR APPLICATION NUMBER: 60/075,409
/ PRIOR FILING DATE: 1998-02-20
/ PRIOR APPLICATION NUMBER: 09/027,287
```

```
/ EARLIER FILING DATE: 1998-02-20
/ EARLIER APPLICATION NUMBER: 09/003,886
/ EARLIER FILING DATE: 1998-01-07
/ EARLIER APPLICATION NUMBER: 08/822,953
/ EARLIER FILING DATE: 1997-03-21
/ EARLIER APPLICATION NUMBER: 60/013,923
/ EARLIER FILING DATE: 1996-03-22
/ EARLIER APPLICATION NUMBER: 60/030,157
/ EARLIER FILING DATE: 1996-10-31
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 1169
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (49)..(768)
US-09-523-323-1

Query Match      70.4% Score 19; DB 4; Length 1169;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 GCGCTCATGAGCAGCAGCAGCCAC 27
DB      717 GCGTTCATCCAGCAGCAGCAGCCAC 691

RESULT 10
US-08-913-014A-6/c
/ Sequence 6, Application US/08913014A
/ Patent No. 6235878
/ GENERAL INFORMATION:
/ APPLICANT: Nishi, Kazunori
/ APPLICANT: Hikiuchi, Yukiko
/ APPLICANT: Shintani, Yasushi
/ TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: David G. Conlin, Esq.
/ ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & USHMAN, LLP
/ STREET: 130 Water Street
/ CITY: Boston,
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/913,014A
/ FILING DATE: 04-SEP-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP97/02480
/ FILING DATE: July 17, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: David G. Conlin
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 342/47694
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1353
/ TYPE: Nucleic acid
/ STRANDEDNESS: Double
```

TOPOLOGY: Linear
MOLECULE TYPE: cDNA
US-08-913-014A-6

Query Match 70.4%; Score 19; DB 3; Length 1353;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCGAC 27
DB 743 GCGTTCATCCAGCAGCAGCAGCCAC 717

RESULT 11
US-09-653-285-6/c

Sequence 6, Application US/09653285
Patent No. 6590090
GENERAL INFORMATION:

APPLICANT: Nishi, Kazumori

Shintani, Yasushi

TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David G. Conlin, Esq.

STREET: 130 Water Street

CITY: Boston,

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/653,285

FILING DATE: 31-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/913,014

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: David G. Conlin

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 342/47694

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1353

TYPE: Nucleic acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-653-285-6

Query Match 70.4%; Score 19; DB 4; Length 1353;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCGAC 27
DB 743 GCGTTCATCCAGCAGCAGCAGCCAC 717

RESULT 12
US-08-913-014A-5/c

Sequence 5, Application US/08913014A

Patent No. 6235878

GENERAL INFORMATION:

APPLICANT: Nishi, Kazumori

Shintani, Yasushi

TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David G. Conlin, Esq.

STREET: 130 Water Street

CITY: Boston,

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/913,014A

FILING DATE: 04-SEP-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP97/02480

FILING DATE: July 17, 1997

ATTORNEY/AGENT INFORMATION:

NAME: David G. Conlin

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 342/47694

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1491

TYPE: Nucleic acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: cDNA

US-08-913-014A-5

Query Match 70.4%; Score 19; DB 3; Length 1491;
Best Local Similarity 81.5%; Pred. No. 26;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCGAC 27
DB 1051 GCGTTCATCCAGCAGCAGCAGCCAC 1025

RESULT 13
US-09-653-285-5/c

Sequence 5, Application US/09653285

Patent No. 6590090

GENERAL INFORMATION:

APPLICANT: Nishi, Kazumori

Shintani, Yasushi

TITLE OF INVENTION: NOVEL FAS LIGAND-LIKE PROTEIN, ITS

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: David G. Conlin, Esq.

STREET: 130 Water Street

CITY: Boston,

STATE: MA

COUNTRY: USA

```
/
/
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/653,285
/ FILING DATE: 31-Aug-2000
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/913,014
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: David G. Conlin
/ REGISTRATION NUMBER: 27,026
/ REFERENCE/DOCKET NUMBER: 342/47694
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1491
/ TYPE: Nucleic acid
/ STRANDEDNESS: Double
/ TOPOLOGY: Linear
/ MOLECULE TYPE: cDNA
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-653-285-5

Query Match          70.4%; Score 19; DB 4; Length 1491;
Best Local Similarity 81.5%; Pred. No. 26;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCGCTATGAGACACGACGCCAC 27
DB 1051 GCGTTCATCCAGCACGACGACAC 1025

RESULT 14
US-09-270-767-12691/c
/ Sequence 12691. Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 12691
/ LENGTH: 1533
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
/ FEATURE:
/ OTHER INFORMATION: n means any nucleotide
US-09-270-767-12691

Query Match          68.9%; Score 18.6; DB 4; Length 1533;
Best Local Similarity 84.0%; Pred. No. 38;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTATGAGACACGACGCCCC 25
DB 734 GTGCTCATGATCACCACGACACC 710
```

```
RESULT 15
US-10-011-200-1
/ Sequence 1. Application US/10011200
/ Patent No. 6632631
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Shuster, Jeffrey R
/ APPLICANT: Tanzer, Matthew M
/ APPLICANT: Hamer, Lisbeth
/ APPLICANT: Adachi, Kiichi
/ APPLICANT: Dezwain, Todd M
/ APPLICANT: Lo, Sze Chung C
/ APPLICANT: Montenegro-Chamorro, Maria V
/ APPLICANT: Darveaux, Blaise A
/ APPLICANT: Frank, Sheryl A
/ APPLICANT: Heiniger, Ryan W
/ APPLICANT: Mahanty, Sanjoy K
/ APPLICANT: Pan, Huaqin
/ APPLICANT: Skelchunes, Amy R
/ APPLICANT: Tarpey, Rex
/ TITLE OF INVENTION: METHODS FOR THE IDENTIFICATION OF INHIBITORS OF HOMOCITRATE SYNTH
/ FILE REFERENCE: 2123US
/ CURRENT APPLICATION NUMBER: US/10/011,200
/ CURRENT FILING DATE: 2001-11-09
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1458
/ TYPE: DNA
/ ORGANISM: Magnaporthe grisea
US-10-011-200-1

Query Match          67.4%; Score 18.2; DB 4; Length 1458;
Best Local Similarity 87.0%; Pred. No. 57;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CTCATGAGACACGACGCCCA 26
DB 484 CTCATGAGACACGACGCCCA 506

Search completed: December 15, 2004, 02:56:25
Job time : 10.2713 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:25:49 : Search time 13.7935 seconds
(without alignments)
10792.424 Million cell updates/sec

Title: US-09-995-225B-41

Sequence: 1 ggcgcacatgagacacacgcacgcacac 27

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	9	US-09-995-225-41
2	27	100.0	27	10	US-09-995-225-41
3	27	100.0	1130	16	US-10-333-946-25
4	27	100.0	1202	14	US-10-094-417-1
5	27	100.0	1526	14	US-10-012-140-10
6	27	100.0	2117	17	US-10-779-104-1
7	27	100.0	2189	14	US-10-219-834-6
8	27	100.0	2188	15	US-10-314-076-1
9	22	81.5	32	16	US-10-296-294A-1
10	22	81.5	42	15	US-10-174-364-117
11	22	81.5	42	16	US-10-246-583-117
12	21	77.8	1038	14	US-10-094-417-19

13	21	77.8	1059	16	US-10-450-590-1
14	21	77.8	1062	9	US-09-995-225-15
15	21	77.8	1062	10	US-09-995-225-15
16	21	77.8	1062	14	US-10-012-140-12
17	21	77.8	1062	16	US-10-296-294A-3
18	21	77.8	1062	16	US-10-296-294A-4
19	21	77.8	1062	16	US-10-450-590-9
20	21	77.8	1110	15	US-10-314-076-3
21	21	77.8	1119	10	US-09-813-432-19
22	21	77.8	1119	15	US-10-174-364-19
23	21	77.8	1119	16	US-10-246-583-19
24	21	77.8	1119	17	US-10-689-832-19
25	21	77.8	1119	17	US-10-689-832-21
26	21	77.8	1119	17	US-10-689-832-21
27	21	77.8	1119	17	US-10-689-832-21
28	21	77.8	1119	17	US-10-689-832-21
29	21	77.8	1119	17	US-10-689-832-21
30	20.2	74.8	1398	16	US-10-425-115-115292
31	19.6	72.6	8838	15	US-10-093-463-81
32	19.6	72.6	2256646	17	US-10-470-555-1
33	19	70.4	612	16	US-10-468-161-30
34	19	70.4	672	18	US-10-425-115-169941
35	19	70.4	720	16	US-10-468-161-4
36	19	70.4	810	13	US-10-066-209-5
37	19	70.4	1017	9	US-09-027-287-38
38	19	70.4	1017	9	US-09-252-656B-38
39	19	70.4	1017	16	US-10-375-680-38
40	19	70.4	1169	9	US-09-027-287-1
41	19	70.4	1169	9	US-09-252-656B-1
42	19	70.4	1169	10	US-09-967-604-5
43	19	70.4	1169	15	US-10-272-411-12
44	19	70.4	1169	15	US-10-218-547-33
45	19	70.4	1169	15	US-10-272-328A-12

ALIGNMENTS

RESULT 1.

US-09-995-225-41

Sequence 41, Application US/09995225

Publication No. US20020193584A1

GENERAL INFORMATION:

APPLICANT: Chen, Ruoping

APPLICANT: Chu, Zhi Liang

APPLICANT: Dang, Huang T.

APPLICANT: Lowitz, Kevin P.

APPLICANT: Pidge, Cameron

TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human

TITLE OF INVENTION: Receptors

FILE REFERENCE: AREN-0308

CURRENT APPLICATION NUMBER: US/09/995,225

CURRENT FILING DATE: 2001-11-26

PRIOR APPLICATION NUMBER: 09/170,496

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: PCT/US99/23938

PRIOR FILING DATE: 1998-10-13

PRIOR APPLICATION NUMBER: 60/253,404

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/255,366

PRIOR FILING DATE: 2000-12-12

PRIOR APPLICATION NUMBER: 60/270,286

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,365

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/270,266

PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/282,032

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,358

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: 60/282,356

PRIOR FILING DATE: 2001-04-06

```

; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020193584A1e1 Sequence
US-09-995-225-41
```

```
Query Match          100.0%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 GCGCTCATGAGCAGCAGCAGCCGCGC 27
Db      1 GCGCTCATGAGCAGCAGCAGCCGCGC 27
```

```
RESULT 2
US-09-995-225-41
; Sequence 41, Application US/09995225
; Publication No. US20030139588A9
```

```

; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huong T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pridle, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A3-Endogenous Versions of Human G
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-41
```

Query Match 100.0%; Score 27; DB 10; Length 27;

```

Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      1 GCGCTCATGAGCAGCAGCAGCCGCGC 27
Db      1 GCGCTCATGAGCAGCAGCAGCCGCGC 27
```

```
RESULT 3
US-10-333-946-25
; Sequence 25, Application US/10333946
; Publication No. US20040023252A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.
; APPLICANT: ARVIZU, Chandra S.; TAL, Preeti G.
; APPLICANT: BURFORD, Neil; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; ELIOT, Vicki S.
; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Nandinder K.
; APPLICANT: HARALTA, April J.A.; YAO, Monique G.
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.
; APPLICANT: POLICKI, Jennifer L.; KEARNEY, Liam
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.
; APPLICANT: LEE, Ernestine A.; DING, Li
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0176 USN
; CURRENT APPLICATION NUMBER: US/10/333,946
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/23433
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/221,478
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/223,268
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/227,054
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/231,121
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,243
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 60/232,691
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/235,146
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7474767CB1
US-10-333-946-25
```

```
Query Match          100.0%; Score 27; DB 16; Length 1130;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 GCGCTCATGAGCAGCAGCAGCCGCGC 27
Db      6 GCGCTCATGAGCAGCAGCAGCCGCGC 32
```

```
RESULT 4
US-10-094-417-1
; Sequence 1, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiegang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
```

Query Match 100.0%; Score 27; DB 10; Length 27;

APPLICANT: Tularik Inc.
TITLE OF INVENTION: No. US20030045685A1el Receptors
FILE REFERENCE: 018781-008110US
CURRENT APPLICATION NUMBER: US/10/094,417
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/802,803
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: US 60/276,649
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1202
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
NAME/KEY: CDS
LOCATION: (68) (1129)
OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
US-10-094-417-1

Query Match 100.0%; Score 27; DB 14; Length 1202;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCGCAC 27
DB 62 GCGCTCATGAGCAGCAGCGCCGCAC 88

RESULT 5
US-10-012-140-10
Sequence 10, Application US/10012140
Publication No. US2003009017A1
GENERAL INFORMATION:
APPLICANT: Leiby, Kevin R.
APPLICANT: Kapeller-Lidemann, Rosana
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
TITLE OF INVENTION: THERBOP
FILE REFERENCE: 381552004900
CURRENT APPLICATION NUMBER: US/10/012,140
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: 60/246,768
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,772
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,185
PRIOR FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 1526
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (139) (1200)
US-10-012-140-10

Query Match 100.0%; Score 27; DB 14; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCGCAC 27
DB 133 GCGCTCATGAGCAGCAGCGCCGCAC 159

RESULT 6

US-10-779-104-1
Sequence 1, Application US/10779104
Publication No. US20040161799A1
GENERAL INFORMATION:
APPLICANT: Andrew J. Murphy
APPLICANT: Susan Croli-Kalish
TITLE OF INVENTION: KOR3L-like-Proteins and Methods of Modulating KOR3L-Mediated Activ
FILE REFERENCE: REG 1000A
CURRENT APPLICATION NUMBER: US/10/779,104
CURRENT FILING DATE: 2004-02-13
PRIOR APPLICATION NUMBER: 60/447,447
PRIOR FILING DATE: 2003-02-14
PRIOR APPLICATION NUMBER: 60/495,577
PRIOR FILING DATE: 2003-08-14
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2117
TYPE: DNA
ORGANISM: homo sapiens
US-10-779-104-1

Query Match 100.0%; Score 27; DB 17; Length 2117;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCGCAC 27
DB 274 GCGCTCATGAGCAGCAGCGCCGCAC 300

RESULT 7
US-10-219-834-6
Sequence 6, Application US/10219834
Publication No. US20030096751A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE TH
FILE REFERENCE: D0191 NE
CURRENT APPLICATION NUMBER: US/10/219,834
CURRENT FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: US 60/313,658
PRIOR FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: US 60/340,703
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: US 60/318,675
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: US 60/355,596
PRIOR FILING DATE: 2002-02-06
PRIOR APPLICATION NUMBER: US 60/333,417
PRIOR FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: US 60/338,367
PRIOR FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 192
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 2189
TYPE: DNA
ORGANISM: Homo sapiens
US-10-219-834-6

Query Match 100.0%; Score 27; DB 14; Length 2189;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCGCCGCAC 27
DB 1041 GCGCTCATGAGCAGCAGCGCCGCAC 1067

RESULT 8
US-10-314-076-1
Sequence 1, Application US/10314076

```
/ Publication No. US20030152977A1
/ GENERAL INFORMATION:
/ APPLICANT: Bristol-Myers Squibb Company
/ TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMT34, AND VARIANTS A
/ FILE REFERENCE: D0197NP
/ CURRENT APPLICATION NUMBER: US/10/314,076
/ PRIOR FILING DATE: 2002-12-06
/ PRIOR APPLICATION NUMBER: U.S. 60/338,371
/ PRIOR FILING DATE: 2001-12-06
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 1
/ LENGTH: 2198
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1047)..(2162)
/ OTHER INFORMATION:
US-10-314-076-1
```

```
Query Match          100.0%; Score 27; DB 15; Length 2198;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GCGCTCATGGAGCACGACGCGCCAC 27
Db      1041 GCGCTCATGGAGCACGACGCGCCAC 1067
```

```
RESULT 9
US-10-296-294A-1
/ Sequence 1, Application US/10296294A
/ Publication No. US20040029224A1
/ GENERAL INFORMATION:
/ APPLICANT: TERAO, Yasuko
/ APPLICANT: MATSUI, Hideki
/ APPLICANT: SHINTANI, Yasushi
/ TITLE OF INVENTION: No. US20040029224A1e1 G Protein-Coupled Receptor and its DNA
/ FILE REFERENCE: 2734 USOP
/ CURRENT APPLICATION NUMBER: US/10/296,294A
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: PCT/JP01/04643
/ PRIOR FILING DATE: 2001-06-01
/ PRIOR APPLICATION NUMBER: JP 2000-170446
/ PRIOR FILING DATE: 2000-06-02
/ PRIOR APPLICATION NUMBER: JP 2000-194926
/ PRIOR FILING DATE: 2000-06-23
/ NUMBER OF SEQ ID NOS: 8
/ SEQ ID NO 1
/ LENGTH: 32
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Designed oligonucleotide primer to amplify DNA encoding TGR8
US-10-296-294A-1
```

```
Query Match          81.5%; Score 22; DB 16; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 CATGGAGCACGACGCGCCAC 27
Db      6 CATGGAGCACGACGCGCCAC 27
```

```
RESULT 10
US-10-174-364-117
/ Sequence 117, Application US/10174364
/ Publication No. US20030216308A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson et al.
```

```
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2
/ CURRENT APPLICATION NUMBER: US/10/174,364
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,836
/ PRIOR FILING DATE: 2000-03-29
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 128
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 117
/ LENGTH: 42
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
US-10-174-364-117
```

```
Query Match          81.5%; Score 22; DB 15; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 CATGGAGCACGACGCGCCAC 27
Db      12 CATGGAGCACGACGCGCCAC 33
```

```
RESULT 11
US-10-246-583-117
/ Sequence 117, Application US/10246583
/ Publication No. US20040058862A1
/ GENERAL INFORMATION:
/ APPLICANT: Majumder
/ TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-729CIP2CON1
/ CURRENT APPLICATION NUMBER: US/10/246,583
/ PRIOR FILING DATE: 2002-12-06
/ PRIOR APPLICATION NUMBER: 10/174,364
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/190,835
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
```


PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 117
LENGTH: 42
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide
US-10-246-583-117

Query Match 81.5%; Score 22; DB 16; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CATGAGCAGCAGCGCCGAC 27
DB 12 CATGAGCAGCAGCGCCGAC 33

RESULT 12
US-10-094-417-19
Sequence 19, Application US/10094417
Publication No. US20030045685A1
GENERAL INFORMATION:
APPLICANT: Tian, Hui
APPLICANT: Zhao, Jiayang
APPLICANT: Chen, Jin-Long
APPLICANT: Cutler, Gene
APPLICANT: Tulek Inc.
TITLE OF INVENTION: NO. US20030045685A1 Receptors
FILE REFERENCE: 018781-008110US
CURRENT APPLICATION NUMBER: US/10/094,417
CURRENT FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/802,803
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: US 60/276,649
PRIOR FILING DATE: 2001-03-16
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 1038
TYPE: DNA
ORGANISM: Mus sp.
FEATURE:
OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR20
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1038)
OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR20
US-10-094-417-19
Query Match 77.8%; Score 21; DB 14; Length 1038;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCGCCGAC 27
DB 1 ATGAGCAGCAGCGCCGAC 21

RESULT 13
US-10-450-590-1
Sequence 1, Application US/10450590
Publication No. US20040076985A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
FILE REFERENCE: LIO316 Foreign Countries
CURRENT APPLICATION NUMBER: US/10/450,590

CURRENT FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/254,923
PRIOR FILING DATE: 2000-12-14
PRIOR APPLICATION NUMBER: US 60/280,110
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: US 60/299,474
PRIOR FILING DATE: 2001-06-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1059
TYPE: DNA
ORGANISM: Homo sapiens
US-10-450-590-1

Query Match 77.8%; Score 21; DB 16; Length 1059;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCAGCAGCGCCGAC 27
DB 1 ATGAGCAGCAGCGCCGAC 21

RESULT 14
US-09-995-225-15
Sequence 15, Application US/09995225
Publication No. US20020193584A1
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huang T.
APPLICANT: Lowitz, Kevin P.
TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
FILE REFERENCE: AREN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 1062
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20020193584A1 Sequence
US-09-995-225-15

Query Match 77.8%; Score 21; DB 9; Length 1062;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCACACGCGCCAC 27
|||||
Db 1 ATGAGCACACGCGCCAC 21

RESULT 15

US-09-995-225-15
; Sequence 15, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Priddy, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human C
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9el Sequence
US-09-995-225-15

Query Match 77.8%; Score 21; DB 10; Length 1062;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ATGAGCACACGCGCCAC 27
|||||
Db 1 ATGAGCACACGCGCCAC 21

Search completed: December 15, 2004, 02:53:12
Job time : 15.7935 secs

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source
Location/Qualifiers

1..358
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G630025017"
/issue_type="spinal cord"
/dev_stage="11 days embryo"
/clone_lib="RIKEN full-length enriched, 11 days embryo spinal cord"

ORIGIN

Query Match 100.0%; Score 27; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGACACGACGCCCCAC 27
|||||
Db 225 GCGCTCATGAGACACGACGCCCCAC 251

RESULT 2 693 bp mRNA linear EST 20-FEB-2001
BF430732 OG04A113 OG Oryza sativa (indica cultivar-group) cDNA clone
LOCUS OG04A113 5' similar to RIBULOSE BIPHOSPHATE CARBOXYLASE SMALL CHAIN
DEFINITION C PRECURSOR, mRNA sequence.
BF430732
BF430732.1 GI:11442833

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 693)
Cushman,J.C. and Nomura,M.
A cDNA library from Oryza sativa cv. Pokkali leaves enriched for
salt-responsive genes
Unpublished (2000)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

JOURNAL
COMMENT
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

FORWARD: T7
BACKWARD: T3
Plate: OG04 row: A column: 11
Seq primer: T3
High quality sequence stop: 350
POLYA=No.

FEATURES
source
Location/Qualifiers

1..693
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="Pokkali"
/db_xref="taxon:39946"
/clone="OG04A113"
/issue_type="leaf"
/dev_stage="7 day-old (treated with 150 mM NaCl for 1, 2 and 7 days)."
/lab_host="Escherichia coli"
/clone_lib="OG"

ORIGIN

Query Match 76.3%; Score 20.6; DB 2; Length 693;
Best Local Similarity 85.2%; Pred. No. 3.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGACACGACGCCCCAC 27
|||||
Db 426 GCGCTCATGAGACACGACGCCCCAC 452

RESULT 3 771 bp DNA linear GSS 16-MAY-2003
CC340173 OGUA629TH ZM 0.7-1.5 KB Zea mays genomic clone ZMBA0392E09,
LOCUS genomic survey sequence.
DEFINITION CC340173
CC340173.1 GI:30809579

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 771)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,W.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: TR
Clas: sheared ends.

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT
Unpublished (2002)
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: TR
Clas: sheared ends.

FEATURES
source
Location/Qualifiers

1..771
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBA0392E09"
/clone_lib="ZM 0.7-1.5 KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 76.3%; Score 20.6; DB 8; Length 771;
Best Local Similarity 85.2%; Pred. No. 3.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGACACGACGCCCCAC 27
|||||
Db 514 GAGCTCTGAGACACGACGCCCCAC 540

RESULT 4 786 bp mRNA linear EST 25-NOV-2002
BU207814/c 603102460F1 CSROCHN03 Gallus gallus cDNA clone CHEST36916 5', mRNA
LOCUS sequence.
ACCESSION BU207814
VERSION BU207814.1 GI:25377508
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Gallus gallus

REFERENCE

1 (bases 1 to 786)
 Phasianinae; Gallus.
 Boardman, P.E., Sant-Esquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 JOURNAL
 Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE
 22355534
 PUBMED
 12445392

COMMENT

Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 016123560409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers

FEATURES

source

```
1..786
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHST36g16"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_1lb="CSBOCHN03"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site_1: EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer. Using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
EcoRI, size-selected, and ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
```

ORIGIN

Query Match 76.3%; Score 20.6; DB 5; Length 786;
 Best Local Similarity 85.2%; Pred. No. 3.8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCCGCGC 27
 |||||
 Db 111 GCGCTCATGAGCAGCAGCCGCGC 85

RESULT 5
 BZ714949 825 bp DNA linear GSS 19-FEB-2003
 LOCUS
 DEFINITION
 BZ714949 genomic survey sequence.
 BZ714949
 VERSION
 BZ714949.1 GI:28435046
 ACCESSION
 BZ714949
 KEYWORDS
 GSS.
 SOURCE
 Zea mays
 ORGANISM
 Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 825)
 White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
 Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
 Benmetzen, J.
 Zea Genomics Consortium
 Unpublished (2003)
 TIGR
 Contact: Cathy Whitelaw

REFERENCE

AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES

source

```
1..825
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8B7a272F20"
/clone_1lb="ZM_0.6_1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
```

ORIGIN

Query Match 76.3%; Score 20.6; DB 8; Length 825;
 Best Local Similarity 85.2%; Pred. No. 3.8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCTCATGAGCAGCAGCCGCGC 27
 |||||
 Db 632 GAGCTGCTGAGCAGCAGCCGCGC 658

RESULT 6

CC396746 109 bp DNA linear GSS 19-MAY-2003
 LOCUS
 DEFINITION
 PHHD211B ZM_0.6_1.0 KB Zea mays genomic clone ZM8B7a29B22,
 genomic survey sequence.

ACCESSION
 CC396746
 VERSION
 CC396746.1 GI:30876836
 KEYWORDS
 GSS.
 SOURCE
 Zea mays
 ORGANISM
 Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 109)
 White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
 Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
 Benmetzen, J.
 Zea Genomics Consortium
 Unpublished (2003)
 TIGR
 Contact: Cathy Whitelaw

TITLE
 JOURNAL
 COMMENT

9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers

FEATURES

source

```
1..109
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM8B7a29B22"
/clone_1lb="ZM_0.6_1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
```

ORIGIN

Query Match 74.8%; Score 20.2; DB 8; Length 109;
 Best Local Similarity 86.0%; Pred. No. 5.3e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GCTCATGAGCAGCAGCCGCGC 27
 |||||
 Db 87 GCTTATGAGCAGCAGCCGCGC 63

RESULT 7
 B0143379/c 539 bp mRNA linear EST 24-APR-2002
 LOCUS fhmhc.p004.e14 Metarhizium anisopliae sf. acridum ARSEF 324
 DEFINITION Metarhizium anisopliae var. acridum cDNA, mRNA sequence.
 ACCESSION B0143379
 VERSION B0143379.1 GI:20280438
 KEYWORDS EST.
 SOURCE Metarhizium anisopliae var. acridum
 ORGANISM Metarhizium anisopliae var. acridum
 Baktayota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Hypocreomycetidae; Hypocreales; Clavicipitaceae; mitosporic
 Clavicipitaceae; Metarhizium.
 REFERENCE 1 (bases 1 to 539)
 AUTHORS Frenosier, F.M., Screen, S., Baga, S., Hu, G. and St. Leger, R.J.
 TITLE EST analysis of genes expressed by two different insect pathogenic
 fungi during optimized secretion of proteins
 JOURNAL Unpublished (2002)
 COMMENT Contact: Frenosier F. M.
 Department of Entomology
 University of Maryland
 4112 Plant Sciences Building, College Park, MD 20742, USA
 Tel: 301 405 16 13
 Fax: 301 314 92 90
 Email: ff34@mail.umd.edu.
 FEATURES
 source
 1. 539
 Location/Qualifiers
 /organism="Metarhizium anisopliae var. acridum"
 /mol_type="mRNA"
 /strain="ARSEF 324"
 /db_xref="taxon:92637"
 /clone_1ib="Metarhizium anisopliae sf. acridum ARSEF 324"
 /note="Vector: Unizap; Metarhizium anisopliae sf. acridum
 was grown on insect cuticle and chitin for 24 hours. A
 cDNA library was constructed in the unidirectional Lambda
 vector Unizap."
 ORIGIN
 Query Match 74.8%; Score 20.2; DB 5; Length 539;
 Best Local Similarity 88.0%; Pred. No. 5.5e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 GCTCATGAGACACGACGCCAC 27
 |||||||
 314 GCTCATGATGCCACGACGCCAC 290
 |||||||
 RESULT 8
 BY324213/c 356 bp mRNA linear EST 11-DEC-2002
 LOCUS BY324213 RIKEN full-length enriched, synovial fibroblasts Mus
 DEFINITION musculus cDNA clone L030023G11 5', mRNA sequence.
 ACCESSION BY324213
 VERSION BY324213
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Baktayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
 1 (bases 1 to 356)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nishida, I., Otsu, N., Saito, R., Suzuki, H., Yamana, A.,
 Miyazawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schombach, C., Gojobori, T., Baldairelli, R., Hill, D.P., Bull, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Busic, V.,
 Chotina, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gietz, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C.A., Secou, M., Shinada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Veirado, R., Wagner, L., Wahlestedt, C., Wang, Y., Warande, Y.,
 Wells, C., Wilming, L.G., Wyshaw-Borja, A., Yamagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shingawa, A., Yasunishi, A., Yoshino, M., Waterson, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 22354683
 PUBLISHED 1246851
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tsgami, M., Waki, K., Watanabe, M., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multichannel sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Vassilie Aidinis (Biomedical Sciences
 Research Center "Al. Fleming" Institute of Immunology 14-16 Al.
 Fleming street 16672 Varti, Greece) whose assistance we gratefully
 acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.
 FEATURES
 source
 1. 356
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="L030023G11"
 /cell_type="synovial fibroblasts"
 /clone_1ib="RIKEN full-length enriched, synovial
 fibroblasts"
 ORIGIN
 Query Match 72.6%; Score 19.6; DB 5; Length 356;
 Best Local Similarity 84.6%; Pred. No. 9.8e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GGCTCATGAGACACGACGCCCA 26

Db 31 GCCGCGACGACGACGACGACGACCA 6

RESULT 9
LOCUS CB388569
DEFINITION CB388569_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
VERSION CB388569.1 GI:30730279
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota: Metazoa: Nematoda: Chromadorea: Rhabditida: Rhabditidae: Rhabditidae: Peloderinae: Caenorhabditis. 1 (bases 1 to 412)
AUTHORS Rebol, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R., Moore, T., Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S., Tolias, P., Pécsek, J., Snyder, M., Papasotiriou, V., Doucette-Stamm, L., Hill, D.E. and Vidal, M.
C. elegans ORFome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression
JOURNAL Nat. Genet. (2003) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFome cloning project: Contact david_hill@dfci.harvard.edu or marc_vidal@dfci.harvard.edu
POLY(A)-No.

FEATURES
source
1. 412
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmcDNA"
/note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauers and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86"

ORIGIN
Query Match 72.6%; Score 19.6; DB 6; Length 412;
Best Local Similarity 84.6%; Pred. No. 9.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCGCTCATGGACGACGACGACCA 26
Db 354 GCGCTCATGGACGACGACGACCA 379

RESULT 10
LOCUS AQ557735/c
DEFINITION AQ557735 455 bp DNA linear GSS 29-MAY-1999
HS_2083_A2_B12_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2083 Col=24 Row=C, genomic survey
sequence.
ACCESSION AQ557735
VERSION AQ557735.1 GI:4917467
KEYWORDS GSS.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo. 1 (bases 1 to 455)
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PubMed 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsec.washington.edu
Plate: 2083 row: C column: 24
Seq primer: T7
Class: BAC ends
High quality sequence stop: 455.

FEATURES
source
1. 455
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=2083 Col=24 Row=C"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelOBAC1; BAC Clones in E-Coli DH10B"

ORIGIN
Query Match 72.6%; Score 19.6; DB 8; Length 455;
Best Local Similarity 84.6%; Pred. No. 9.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CGGTCGTAGACACCGACGCCCCAC 27
Db 323 CGGTCGTAGACACCGACGCCCCAC 298

RESULT 11
LOCUS AQ177428
DEFINITION AQ177428 490 bp DNA linear GSS 17-OCT-1998
HS_2206_A2_F06 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2206 Col=12 Row=K, genomic survey
sequence.
ACCESSION AQ177428
VERSION AQ177428.1 GI:3574795
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo. 1 (bases 1 to 490)
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PubMed 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2206 Row: K Column: 12
 Class: BAC ends
 High quality sequence stop: 490.

FEATURES

source

1..490
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=2206 Col=12 Row=K"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelBAC11; BAC clones in E-coli DH10B"

ORIGIN

Query Match 72.6%; Score 19.6; DB 8; Length 490;
 Best Local Similarity 84.6%; Pred. No. 9.9e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CGCTATGAGACACGACGCCAC 27
 |||||
 Db 428 CCCCATGAGACACTCTCCAC 453

RESULT 12
 A2446249 504 bp DNA linear GSS 04-OCT-2000
 LOCUS 1M0242A10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

ACCESSION A2446249
 VERSION A2446249.1 GI:10596873
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 504)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmood, M., Meenen, E., Pedersen, T.,
 Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
 Niederhausern, A. and Wright, D., Weis, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
 Plate: 0242 row: A column: 10
 Seq primer: CACACGAGAAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 504.

FEATURES

source

1..504
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0242A10"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PMD42nv, Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/nares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g1|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

Query Match 72.6%; Score 19.6; DB 8; Length 504;
 Best Local Similarity 84.6%; Pred. No. 9.9e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CGCTATGAGACACGACGCCAC 27
 |||||
 Db 341 CGCATGACGACACACGCGCAC 366

RESULT 13
 AM057282 599 bp mRNA linear EST 29-SEP-1999
 LOCUS ca03c06.y1 C elegans fem3 Q23 S1 Caenorhabditis elegans cDNA 5'

ACCESSION AM057282
 VERSION AM057282.1 GI:5932921
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans

ORGANISM

Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditidae; Rhabditidae; Pelodermidae; Caenorhabditis.

REFERENCE

1 (bases 1 to 599)
 Ward, S., Smith, H., Clifton, S., Marra, M., Hillier, L., Kucaba, T.,
 Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M.,
 Harvey, N., Ritzer, E., Jackson, Y., McCann, R., Waterston, R. and
 Wilson, R.

JOURNAL Unpublished (1999)
 COMMENT Contact: Samuel Ward, Ph.D.
 UofArizona-Washu C. elegans EST project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 Contact Harold Smith (hes@u.arizona.edu) for further information
 relating to organism, libraries, or clone availability.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 486.

FEATURES

source

1..599
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /db_xref="taxon:6239"
 /lab_host="DH5alpha cells"
 /clone_lib="C elegans fem3 Q23 S1"
 /note="Vector: pluescript II SK+, Site 1: XhoI; Site 2:
 NotI; This C elegans library was made from fem-3(q23)-
 worms (produce only sperm at 25 C). cDNA was generated
 via oligo (dT) priming and directionally cloned into
 pluescript II SK+ vector with a modified polylinker.
 This library was subtracted with a second fem-1(hc17)
 (produce only oocytes at 25 C) library to enrich this

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:34:50 ; Search time 84.2037 Seconds
(without alignments)
14040.270 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25

Sequence: 1 gaggcagtagtgcacacactatg 25

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing filter 45 summaries

Database : GenEmbl:*
1: gb ha:*
2: gb htg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	AX498219	AX498219 Sequence
2	25	100.0	1343	AX254977	AX254977 Sequence
3	25	100.0	1536	AX664701	AX664701 Sequence
4	25	100.0	1826	AX463232	AX463232 Sequence
5	25	100.0	101882	AC021089	AC021089 Homo sapi
6	25	100.0	110000	AL139235	AL139235 Homo sapi
7	20.2	80.8	186606	AC129308	AC129308 Mus muscu
8	20	80.0	204798	AC125810	AC125810 Rattus no
9	20	80.0	236921	AC094151	AC094151 Rattus no
10	20	80.0	241111	AC118438	AC118438 Rattus no
11	20	80.0	245855	AC106203	AC106203 Rattus no
12	19.2	76.8	142525	AL671501	AL671501 Mouse DNA
13	19.2	76.8	149988	AL591602	AL591602 Human DNA
14	19.2	76.8	194457	AL831765	AL831765 Mouse DNA
15	19.2	76.8	219990	AC112993	AC112993 Mus muscu
16	19.2	76.8	249841	AC095767	AC095767 Rattus no
17	19.2	76.8	257767	AC095974	AC095974 Rattus no
18	19.2	76.8	326663	AF011889	AF011889 Human Xq2
19	19	76.0	1596	AF298642	AF298642 Avian pne

C 20	19	76.0	1596	14	AF298643	AF298643 Avian pne
C 21	19	76.0	1596	14	AF298644	AF298644 Avian pne
C 22	19	76.0	1596	14	AF298645	AF298645 Avian pne
C 23	19	76.0	1596	14	AF298646	AF298646 Avian pne
C 24	19	76.0	1596	14	AF298647	AF298647 Avian pne
C 25	19	76.0	1596	14	AF298648	AF298648 Avian pne
C 26	19	76.0	1596	14	AF298649	AF298649 Avian pne
C 27	19	76.0	1596	14	AF298650	AF298650 Avian pne
C 28	18.8	75.2	11240	9	AC127902	AC127902 Homo sapi
C 29	18.8	75.2	50156	9	AL358053	AL358053 Human DNA
C 30	18.8	75.2	118632	9	AL133380	AL133380 Human DNA
C 31	18.8	75.2	156095	9	AC096577	AC096577 Homo sapi
C 32	18.8	75.2	170313	10	AC121973	AC121973 Mus muscu
C 33	18.8	75.2	176092	2	AC073252	AC073252 Homo sapi
C 34	18.8	75.2	181144	9	AL365444	AL365444 Human DNA
C 35	18.8	75.2	181717	9	AL353696	AL353696 Human DNA
C 36	18.8	75.2	209428	2	AC125948	AC125948 Rattus no
C 37	18.8	75.2	210321	2	AC132027	AC132027 Rattus no
C 38	18.6	74.4	86290	9	AC112503	AC112503 Homo sapi
C 39	18.6	74.4	110000	1	BX950851	Continuation (26 o
C 40	18.6	74.4	142488	10	AC087040	AC087040 Mus muscu
C 41	18.6	74.4	209119	2	AC083818	AC083818 Mus muscu
C 42	18.4	73.6	167123	10	AL663041	AL663041 Mouse DNA
C 43	18.4	73.6	208880	2	AC128103	AC128103 Rattus no
C 44	18.4	73.6	212178	2	AC125925	AC125925 Rattus no
C 45	18.4	73.6	220407	2	AC121626	AC121626 Rattus no

ALIGNMENTS

RESULT 1	AX498219	25 bp	DNA	linear	PAT 26-SEP-2002
LOCUS	AX498219	Sequence 42 from Patent WO0242461.			
DEFINITION	AX498219				
ACCESSION	AX498219.1	GI:23343141			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.				
TITLE	Endogenous and non-endogenous versions of human g protein-coupled				
JOURNAL	receptors				
FEATURES	Patent: WO 0242461-A 42 30-MAY-2002;				
source	Arena Pharmaceuticals, Inc. (US)				
	Location/Qualifiers				
	1..25				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="Novel Sequence"				
ORIGIN					
Query Match	100.0%;	Score 25;	DB 6;	Length 25;	
Best Local Similarity	100.0%;	Pred. No. 0.25;			
Matches	25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1				
Db	1				
	1	GAGGCAGTAGTTCACACACTATG 25			
RESULT 2	AX254977	1343 bp	DNA	linear	PAT 10-OCT-2001
LOCUS	AX254977	Sequence 21 from Patent WO0170978.			
DEFINITION	AX254977				
ACCESSION	AX254977.1	GI:16074504			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens (human)				
	Homo sapiens				

REFERENCE 1
AUTHORS Taupier, R. J., Majumder, K., Spaderna, S. K., Smithson, G., Mezes, P. S.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 01/0978-A 21 27-SEP-2001;
Curagen Corporation (US)
FEATURES
source 1. 1343
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCAGTAGTGGCCACACCTATGG 25
Db 1090 GAGGCAGTAGTGGCCACACCTATGG 1066
RESULT 3
AX664701/c AX664701 1526 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 10 from Patent WO02074960.
DEFINITION AX664701
ACCESSION AX664701
VERSION AX664701.1 GI:29164461
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Leiby, K. R., Kapeller, J., Libermann, R. and Glucksmann, M.
TITLE 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
JOURNAL Patent: WO 02074960-A 10 26-SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1. 1526
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
139. 1200
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD80046.1"
/db_xref="GI:29164462"
/translation="MEHTAHLANSSLSWSPGACGLGFVPTVYISLIQICLPAN
ILTVILISQVARROKSSVNYVLALAAADLVLFVVDLFLEDFILNNQVBPDK
IIIVLEESSIHSTWITVPLTIDRYLAVCPLEKTVSPYARFKYIVSYITCFULTS
IPYWMNINWTEDEYISVSHVHLIWHVCFVGVQVPSIFILMSIVYKLRSSNRL
CGYSGTKTATLFTISIFAILWAPLIMLTHLXGAPIONRPLVHIMDIAMMLALL
NTAINFLYFISKRRTMAATLKAFKCKQKQVQPTTNHNSITSSPWISPSNSHC
IKMLVYQYDNGKPKIKVSP"
ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCAGTAGTGGCCACACCTATGG 25
Db 1227 GAGGCAGTAGTGGCCACACCTATGG 1203
RESULT 4
AX463232/c AX463232 1826 bp DNA linear PAT 15-JUN-2002
LOCUS Sequence 6 from Patent WO0248358.
DEFINITION

ACCESSION AX463232
VERSION AX463232.1 GI:21886202
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Oikigami, H.
TITLE Regulation of human chemokine-like receptor
JOURNAL Patent: WO 0248358-A 6 20-JUN-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1. 1826
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 25; DB 6; Length 1826;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGCAGTAGTGGCCACACCTATGG 25
Db 1132 GAGGCAGTAGTGGCCACACCTATGG 1108
RESULT 5
AC021089 AC021089 101882 bp DNA linear PRI 22-MAR-2003
LOCUS Homo sapiens chromosome 16 clone CTD-2264D9, complete sequence.
DEFINITION AC021089
ACCESSION AC021089
VERSION AC021089.8 GI:29150340
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Unpublished
2 (bases 1 to 101882)
REFERENCE 2
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 101882)
REFERENCE 3
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 101882)
REFERENCE 4
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
5 (bases 1 to 101882)
REFERENCE 5
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Mar 22, 2003 this sequence version replaced gi:20330797.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.

FEATURES
source
1.101882 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2264D9"

ORIGIN
Query Match 100.0%; Score 25; DB 9; Length 101882;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGTACTTTCACACCTATCG 25
18207 GAGCAGTACTTTCACACCTATCG 18231

RESULT 6
AL139235.0
WPCOMMENT
Sequence split into 4 fragments LOCUS AL139235 Accession AL139235
Fragment Name Begin End
AL139235.0 1 110000
AL139235_1 100001 210000
AL139235_2 200001 310000
AL139235_3 300001 407025
AL139235 407025 407025 bp DNA linear HTG 16-JUL-2001
Homo sapiens chromosome 9 clone RP11-3215, 31 unordered pieces.
DEFINITION
AL139235
ACCESSION
AL139235.22 GI:11034497
VERSION
HTG; HTGS_PHASE1; HTGS_CANCELLED.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Plumb, B.
Direct Submission
Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
request: clonerequest@sanger.ac.uk
On Oct 26, 2000 this sequence version replaced gi:11024977.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: DA3215
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 389894 bases at least Q40
Consensus quality: 395933 bases at least Q30
Consensus quality: 400050 bases at least Q20
Insert size: 174695; 17.7% error; agarose-fp
Quality coverage: 4.93x in Q20 bases; sum-of-contigs Quality
coverage: 14.49x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2425 contig of 2425 bp in length
* 2426 2525: gap of 100 bp
* 2526 10023: contig of 7498 bp in length

FEATURES
source
1.10124 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-3215"
/clone_11b="RPC1-11.1"
1.2425
/note="assembly fragment: 05385
fragment_chain: 1"
2526.10023
/note="assembly fragment: 03131
fragment_chain: 1"
10124.36364

10024 10123: gap of 100 bp
10124 36364: contig of 26241 bp in length
* 36365 36454: gap of 100 bp
* 36465 43873: contig of 7409 bp in length
* 43874 43973: gap of 100 bp
* 43974 49988: contig of 6015 bp in length
* 49989 50088: gap of 100 bp
* 50089 93603: contig of 43515 bp in length
* 93604 93703: gap of 100 bp
* 93704 96191: contig of 2488 bp in length
* 96192 96291: gap of 100 bp
* 96292 99433: contig of 3142 bp in length
* 99434 99533: gap of 100 bp
* 99534 101556: contig of 2023 bp in length
* 101557 101657: gap of 100 bp
* 101657 184109: contig of 82453 bp in length
* 184110 184209: gap of 100 bp
* 184210 290696: contig of 106487 bp in length
* 290697 290796: gap of 100 bp
* 290797 292892: contig of 2096 bp in length
* 292893 292992: gap of 100 bp
* 292993 296019: contig of 3026 bp in length
* 296019 322154: contig of 26036 bp in length
* 322155 322254: gap of 100 bp
* 322255 324646: contig of 2392 bp in length
* 324647 324746: gap of 100 bp
* 324747 329069: contig of 4323 bp in length
* 329070 329170: gap of 100 bp
* 329171 331371: contig of 2202 bp in length
* 331372 331471: gap of 100 bp
* 331472 333909: contig of 2438 bp in length
* 333910 334009: gap of 100 bp
* 334010 369445: contig of 35436 bp in length
* 369446 371996: contig of 2451 bp in length
* 371997 372096: gap of 100 bp
* 372097 375925: contig of 3829 bp in length
* 375926 376025: gap of 100 bp
* 376026 378132: contig of 2107 bp in length
* 378133 378232: gap of 100 bp
* 378233 381590: contig of 3358 bp in length
* 381591 381690: gap of 100 bp
* 381691 383726: contig of 2036 bp in length
* 383727 383826: gap of 100 bp
* 383827 390832: contig of 7006 bp in length
* 390833 390932: gap of 100 bp
* 390933 394216: contig of 3284 bp in length
* 394217 394316: gap of 100 bp
* 394317 396358: contig of 2042 bp in length
* 396359 396458: gap of 100 bp
* 396459 398733: contig of 2275 bp in length
* 398734 398833: gap of 100 bp
* 398834 402385: contig of 3552 bp in length
* 402386 402485: gap of 100 bp
* 402486 404522: contig of 2037 bp in length
* 404523 404622: gap of 100 bp
* 404623 407025: contig of 2403 bp in length.

misc_feature
1.2425
misc_feature
2526
misc_feature
10124.36364

```

misc_feature /note="assembly_fragment:06402
fragment_chain:1"
36465..43873
/note="assembly_fragment:01436
fragment_chain:1"
43974..49988
/note="assembly_fragment:04547
fragment_chain:1"
50089..53603
/note="assembly_fragment:01636
fragment_chain:1"
93704..96191
/note="assembly_fragment:02957
fragment_chain:2"
96292..99433
/note="assembly_fragment:00499
fragment_chain:2"
99534..101556
/note="assembly_fragment:03112
fragment_chain:2"
101657..184109
/note="assembly_fragment:01419
fragment_chain:3"
184210..290596
/note="assembly_fragment:00925
fragment_chain:3"
290797..292892
/note="assembly_fragment:03065
fragment_chain:4"
292993..296018
/note="assembly_fragment:05057
fragment_chain:4"
296119..322154
/note="assembly_fragment:00269"
322255..324646
/note="assembly_fragment:00666"
324747..329069
/note="assembly_fragment:01113"
329170..331371
/note="assembly_fragment:01461"
331472..333909
/note="assembly_fragment:01516"
334010..369445
/note="assembly_fragment:01942.0"
369546..371996
/note="assembly_fragment:02194"
372097..375925
/note="assembly_fragment:02249"
376026..378132
/note="assembly_fragment:02485"
378233..381590
/note="assembly_fragment:02649"
381691..383726
/note="assembly_fragment:02869"
383827..390832
/note="assembly_fragment:03540"
390933..394216
/note="assembly_fragment:04405"
394317..396358
/note="assembly_fragment:04595"
396459..398733
/note="assembly_fragment:05768"
398834..402385
/note="assembly_fragment:06106"
402486..404522
/note="assembly_fragment:06204"
404623..407025
/note="assembly_fragment:06292"

```

ORIGIN

```

Query Match      100.0%; Score 25; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GAGCAGTAGTTCACACCTATGG 25
Db      34226 GAGCAGTAGTTCACACCTATGG 34250

RESULT 7
AC129308
LOCUS   AC129308
DEFINITION Mus musculus BAC clone RP24-308E2 from chromosome 14, complete
sequence.
AC129308
VERSION AC129308.6 GI:33285230
KEYWORDS HTG.
SOURCE  Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 186606)
Wang, C., Abbot, A. and Haglund, K.
The sequence of Mus musculus BAC clone RP24-308E2
2 (bases 1 to 186606)
Wilson, R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 186606)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (28-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 186606)
Wilson, R.K.
Direct Submission
Submitted (26-JUL-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 186606)
Wilson, R.K.
Direct Submission
Submitted (27-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 26, 2003 this sequence version replaced gi:31581805.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0308E02
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

```

MAPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see

http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (http://www.chori.org) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
Location/Qualifiers

```

1..186606
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="14"
/map="14"
/clone="RP24-308E2"
/clone_1lb="RPCI-24"
221..574
/repeat_region
/rpc_family="MaLR"
600..715
/repeat_region
/rpc_family="Alu"
1550..1674
/repeat_region
/rpc_family="B4"
2438..3434
/repeat_region
/rpc_family="L1"
3575..3932
/repeat_region
/rpc_family="MaLR"
4945..5054
/misc_feature
5884..7189
/repeat_region
/note="Sequence derived from PCR product of project DNA."
6134..6199
/unsure
/rpc_family="L1"
6268..6377
/unsure
/note="Sequence derived from one plasmid subclone."
7705..8093
/repeat_region
/rpc_family="ERVY"
8952..9065
/repeat_region
/rpc_family="ERV1"
9997..10194
/repeat_region
/rpc_family="B4"
10232..10309
/repeat_region
/rpc_family="ID"
12635..12695
/repeat_region
/rpc_family="ID"
13111..13466
/repeat_region
/rpc_family="MaLR"
14142..14561
/repeat_region
/rpc_family="MaLR"
14566..14764
/repeat_region
/rpc_family="L2"
15512..15615
/repeat_region
/rpc_family="B4"
15616..15770
/repeat_region
/rpc_family="MaLR"
15772..16030
/repeat_region
/rpc_family="L1"
16031..16100
/repeat_region
/rpc_family="B2"
16073..16177
/repeat_region
/rpc_family="B2"
16845..17117
/repeat_region
/rpc_family="L1"
25187..25478
/repeat_region
/rpc_family="MaLR"
25573..25643
/repeat_region
/rpc_family="MaLR"
25929..26013
/repeat_region
/rpc_family="L1"
26476..26521
/repeat_region
/rpc_family="B4"

```

```

repeat_region 26549..26672
/rpc_family="B4"
repeat_region 26805..27276
/rpc_family="MaLR"
repeat_region 27341..27421
/rpc_family="B4"
repeat_region 28427..28713
/rpc_family="MaLR"
repeat_region 29106..29587
/rpc_family="ERV1"
repeat_region 29608..30271
/rpc_family="L1"
repeat_region 32786..33794
/rpc_family="L1"
repeat_region 33850..35375
/rpc_family="L1"
repeat_region 35376..36321
/rpc_family="RMR1A"
repeat_region 36322..38311
/rpc_family="L1"
repeat_region 38372..39557
/rpc_family="L1"
/rpc_family="L1"
39558..39870
/unsure
/note="Unresolved simple sequence repeat."
40130..40209
/rpc_family="Alu"
40194..40421
/rpc_family="MaLR"
41371..41425
/rpc_family="ERV1"
42893..43083
/rpc_family="B2"
46442..46597
/rpc_family="B2"
46738..47109
/rpc_family="B2"
47267..47414
/rpc_family="MaLR"
48365..48497
/rpc_family="Alu"
49562..49772
/rpc_family="B4"
51468..51582
/rpc_family="B2"
54385..54626
/rpc_family="Alu"
55038..55103
/rpc_family="RMR15"
55595..55743
/rpc_family="ERV1"
55909..55998
/rpc_family="Achobo"
56001..56182
/rpc_family="Alu"
57800..57881
/rpc_family="MER1_type"
59721..59939
/rpc_family="MaLR"
60010..60335
/rpc_family="B2"
61776..61870
/rpc_family="RICKSHA"
/rpc_family="MaLR"

```

Query Match 80.8%; Score 20.2; DB 10; Length 186606;
Best Local Similarity 88.0%; Pred. No. 32;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGGCACTAGTTCACCACTATGG 25
DB 11111 GAGGCACTAGTTCACCACTATGG 11135

RESULT 8

AC125810/c
 LOCUS 204798 bp DNA linear HTG 20-NOV-2002
 DEFINITION Rattus norvegicus clone CH230-345P21, WORKING DRAFT SEQUENCE.
 AC125810
 AC125810.3 GI:25139743
 HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 204798)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, S., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhmed, F.,
 Bismalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavares, I., Caesar, H., Center, A.,
 Chacko, U., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Bayes, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregiorgis, E., Geer, K., Gall, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, J.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulys, S., Hume, J., Idelbrad, D., Jackson, A.,
 Jackson, L., Jacob, U., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, D.,
 Lorenshuwa, L., Loulsegod, H., Lozado, R.J., Lu, X., Ma, J.,
 Maheswari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 McWhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaoketeneh, O., Okwomu, G., Olanunwasan, A., Pal, S., Parks, K.,
 Patel, N., Paul, H., Perez, A., Perez, L., Plimkock, C.,
 Plopper, F., Poldexter, A., Popov, C.D., Primus, E., Pu, L., L.,
 Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R.,
 Rell, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, D.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, U., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weiss, R.A., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 204798)
 Worley, K.C.
 Direct Submission
 Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 204798)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department

COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:23915433.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KXDO

Center clone name: CH230-345P21

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 190846 bases at least Q40

Consensus quality: 192519 bases at least Q30

Consensus quality: 193364 bases at least Q20

Estimated insert size: 194659; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length

----- (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

----- NOTE: This is a 'working draft' sequence. It currently

----- consists of 1 contigs. Gaps between the contigs

----- are represented as runs of N. The order of the pieces

----- is believed to be correct as given, however the sizes

----- of the gaps between them are based on estimates that have

----- provided by the submitter.

----- This sequence will be replaced

----- by the finished sequence as soon as it is available and

----- the accession number will be preserved.

----- 1 204798: contig of 204798 bp in length.

----- Location/Qualifiers

1. 204798

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-345P21"

886. 1703

/note="clone_boundary"

clone_end:T7

site:

end_sequence: B2131587

/note="complement(196726..197574)"

/complement(196726..197574)

clone_end: Sp6

site:

end_sequence: B2131588

203715..204798

/note="wgs_end_extension"

clone_end: Sp6

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN
 Query Match 80.0%; Score 20; DB 2; Length 204798;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0;
 Gaps 0;
 2 AGCAGTATTGTCACACCT 21
 Db 70503 AGCAGTATTGTCACACCT 70484

RESULT 9
AC094151/c
LOCUS DEFINITION
AC094151
AC094151.8 GI:30467870
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AC094151 236921 bp DNA linear HTG 09-MAY-2003
Rattus norvegicus clone CH230-2L21, WORKING DRAFT SEQUENCE, 3
unordered pieces.
AC094151
HTG, HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILLOR.
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 236921)
Wuzny, D., Marten, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooke, S., Amin, A., Angiano, D.,
Ayala-Bech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Blewett, K., Blay, J., Blankenburg, K., Bluth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chachyo, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Devilla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Ebeco, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabler, A., Ganta, R., Garcia, A., Garner, T., Gava, M.,
Gibson, G., Gier, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
Gunnarsson, P., Healand, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howell, S., Hulik, S., Hume, J., Idler, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kow, J., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenz, W., L., Louised, H., Lozada, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Mlosov, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nasekelen, O., Okunolu, G., Olariu, A., Pal, S., Parks, K.,
Patek, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.,
Pizarro, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K.,
Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
Williams, G., Willson, R., Wlezyk, R., Woodson, H., Wozniak, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G., and Gibbs, R.A.
Direct Submission
2 (bases 1 to 236921)
Unpublished
Morley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 236921)
Rat Genome Sequencing Consortium.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24819635.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAB2
Center clone name: CH230-2L21
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 228373 bases at least Q40
Consensus quality: 229687 bases at least Q30
Consensus quality: 230766 bases at least Q20
Estimated insert size: 239162; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 232260: contig of 232260 bp in length
* 232261 232360: gap of unknown length
* 232361 233968: contig of 1608 bp in length
* 233969 234068: gap of unknown length
* 234069 236921: contig of 2853 bp in length.
Location/Qualifiers
1. 236921
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-2L21"
complement(531..1218)
/note="clone boundary"
clone_end:17
site:ECORI
end_sequence:BH288928"
complement(231283..231561)
/note="clone boundary"
clone_end:5p6
site:ECORI
end_sequence:BH288930"
ORIGIN
Query Match 80.0%; Score 20; DB 2; Length 236921;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Qy 2 AGCAGTAGTGGCACCCT 21
|||||
Db 189899 AGCAGTAGTGGCACCCT 189880

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AGGAGTAGTGGCAGACT 21
 |||||
 Db 20325 AGGAGTAGTGGCAGACT 203344

RESULT 11
 AC106203
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-27C8, *** SEQUENCING IN PROGRESS ***
 AC106203 245855 bp DNA linear HTG 13-MAY-2003
 AC106203 8 unordered pieces.
 AC106203
 VERSION HTG: HTGS PHASE1; HTGS DRAFT: HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 245855)
 AUTHORS Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., A., Faller, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., Gabler, A., Gante, R., Garcia, A., Garner, T., Gava, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, A., Guerra, M., Guevara, W., Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Hayak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idiebird, D., Jackson, A., Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowic, C., Kraft, C., L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuwa, L., Loulseged, H., Lozada, R., J., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M., P., McNeill, T., Z., Meenen, E., Mlosoajevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munida, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okunolu, G., Olarinmase, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., L., Pizzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M., A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richard, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C., D., Smales, D., Speed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Yoon, L., Yoon, V., Niederhausen, A., Weis, R., Smith, D., R., Holt, R., A., Smith, H., O., Weinstock, G., and Gibbs, R., A.

TITLE JOURNAL
 REFERENCE 2 (bases 1 to 245855)
 Unpublished

AUTHORS Worley, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 245855)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On May 13, 2003 this sequence version replaced gi:22269037. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GJVB
 Center clone name: CH230-27C8
 ----- Summary Statistics
 Assembly program: Atlas 3.0:
 Consensus quality: 212523 bases at least Q40
 Consensus quality: 217649 bases at least Q30
 Consensus quality: 221314 bases at least Q20
 Estimated insert size: 22736; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 32679: contig of 32679 bp in length
 32680 32779: gap of unknown length
 32780 54985: contig of 22206 bp in length
 54986 55085: gap of unknown length
 55086 73449: contig of 18344 bp in length
 73450 73549: gap of unknown length
 73550 211048: contig of 137499 bp in length
 211049 211148: gap of unknown length
 211149 221087: contig of 9939 bp in length
 221088 221187: gap of unknown length
 221188 239949: contig of 18762 bp in length
 239950 240049: gap of unknown length
 240050 241182: contig of 1133 bp in length
 241183 241282: gap of unknown length
 241283 245855: contig of 4573 bp in length.

Location/Qualifiers
 1. 245855
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-27C8"
 73550. 74985
 /note="wgs_contig"

FEATURES
 source misc_feature

misc_feature 75386..77137
/note="wgs_config"
misc_feature 22188..222913
/note="wgs_contig"

ORIGIN

Query Match 80.0%; Score 20; DB 2; Length 24585;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AGGCAGTAGTGGCCACCT 21
|||||
Db 138907 AGGCAGTAGTGGCCACCT 138926

RESULT 12
AL671501 142525 bp DNA 1linear ROD 01-JUL-2003
LOCUS Mouse DNA sequence from clone RP23-313P3 on chromosome X, complete
DEFINITION sequence.
ACCESSION AL671501
VERSION AL671501.8 GI:32400078
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 142525)
Whitehead, S.
Direct Submission
Submitted (01-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Jul 1, 2003 this sequence version replaced gi:3216575.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

COMMENT

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPI-23 Mouse BAC library
constructed by the group of Pletier de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
Location/Qualifiers
1..142525
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"

FEATURES
source

/clone="RP23-313P3"
/clone_1ib="RPCT-23"

ORIGIN

Query Match 76.8%; Score 19.2; DB 10; Length 142525;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGCAGTAGTGGCCACCTATG 25
|||||
Db 103084 AGGCAGTAGTGGCCACCTATG 103107

RESULT 13
AL591602 149988 bp DNA 1linear PRI 16-DEC-2001
LOCUS Human DNA sequence from clone RP5-1092E23 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL591602
VERSION AL591602.7 GI:17907262
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 149988)
Wallis, J.
Direct Submission
Submitted (16-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Dec 18, 2001 this sequence version replaced gi:16973109.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone configs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP5-1092E23 is from the library RPI-5 constructed by the group of
Pletier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP5-1092E23 it may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP5-1092E23 is at 149988 in this
sequence. The true right end of clone RP4-573J3 is at 2000 in this
sequence.

FEATURES
source
1..149988
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP5-1092E23"
/clone_1ib="RPI-5"

ORIGIN

Query Match 76.8%; Score 19.2; DB 9; Length 149988;
 Best Local Similarity 87.5%; Pred. No. 1e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 AGGAGTAGTTGCCACACTATG 25
 |||||
 Db 15256 AGGAGTAGTTGCCACCTCTAG 15279

RESULT 14
 AL831765 194457 bp DNA linear ROD 24-SEP-2002
 LOCUS Mouse DNA sequence from clone RP23-148012 on chromosome 4, complete
 DEFINITION
 accession AL831765
 version AL831765.3 GI:23320939
 keywords HTG.
 source Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Martin, S.
 TITLE Direct Submission
 JOURNAL Submitted (24-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk
 On Sep 25, 2002 this sequence version replaced gi:22798457.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SwissProt; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-148012 is constructed by the group of Pieter de Jong
 For further details see http://www.chori.org/bacpac/home.htm
 VECTOR: pBac3.6.

FEATURES
 source
 1. 194457
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clone="RP23-148012"
 /clone_1fb="RP23-148012"

ORIGIN

Query Match 76.8%; Score 19.2; DB 10; Length 194457;
 Best Local Similarity 87.5%; Pred. No. 1e+02;
 Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 AGGAGTAGTTGCCACACTATG 25
 |||||

Db 155036 AGGAGTAGTTGCCACACTATG 155059

RESULT 15
 AC112943/c 219990 bp DNA linear ROD 21-OCT-2003
 LOCUS Mus musculus chromosome 16, clone RP23-424C7, complete sequence.
 DEFINITION
 accession AC112943
 version AC112943.12 GI:37777390
 keywords HTG.
 source Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Birren, B., Nussbaum, C. and Lander, E.
 TITLE Unpublished
 JOURNAL
 REFERENCE
 AUTHORS
 2 (bases 1 to 219990)

Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Baetien, V., Boguski, L., Boukhalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehoczek, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, R., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menues, L., Mihova, T., Mienna, P., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 JOURNAL Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 219990)
 Birren, B., Nussbaum, C. and Lander, E.
 REFERENCE
 AUTHORS
 JOURNAL
 REFERENCE
 AUTHORS
 Direct Submission
 Submitted (04-SEP-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 219990)

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Baetien, V., Bloom, T., Boguski, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarty, C., McDonald, P., X. Liu, G., Mbbilit, R., Maclean, C., Meldrim, J., Menues, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, C., Retta, R., Riese, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (21-OCT-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Oct 21, 2003 this sequence version replaced gi:34447307.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997).
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

----- Project Information
Contact: sequence.submissions@genome.wi.mit.edu

Center project name: L20313
Center clone name: 424_C_7

FEATURES

source

----- Location/Qualifiers

1. .219990

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="16"

/map="16"

/clone="RP23-424C7"

misc_feature

1. .15027
/note="wgs_end_extension
clone_end:SP6"

repeat_region

1. .1056
/rpt_family="Lx6"

repeat_region

1083. .1621
/rpt_family="Lx6"

repeat_region

1622. .1654
/rpt_family="A-rich"

repeat_region

1655. .2272
/rpt_family="Lx6"

repeat_region

2865. .2966
/rpt_family="PB1D10"

repeat_region

3403. .3542
/rpt_family="B1_MM"

repeat_region

3545. .3681
/rpt_family="GAAAA)n"

repeat_region

3682. .3800
/rpt_family="PB1D9"

repeat_region

5146. .5200
/rpt_family="GA-rich"

repeat_region

complement(5481. .5561)
/rpt_family="MER5A"

repeat_region

complement(6134. .6274)
/rpt_family="B1_MM"

repeat_region

6765. .6784
/rpt_family="CAA)n"

repeat_region

6794. .6799
/note="<30 qual SNGL region"

repeat_region

6894. .6900
/note="<30 qual SNGL region"

repeat_region

6908. .6921
/note="<30 qual SNGL region"

repeat_region

6931. .6943
/note="<20 qual SNGL region"

repeat_region

6939. .6966
/note="<20 qual SNGL region"

repeat_region

7021. .7251
/note="single clone coverage"

repeat_region

7376. .7396
/rpt_family="AT-rich"

repeat_region

7442. .7638
/rpt_family="B3"

repeat_region

8460. .8607
/rpt_family="B1_MM"

repeat_region

8789. .8997
/rpt_family="ORR1D"

repeat_region

9005. .9199
/rpt_family="Lx7"

repeat_region

10246. .10536
/rpt_family="B4"

repeat_region

10727. .10875
/rpt_family="B1_MM"

repeat_region

11642. .11693
/rpt_family="AT-rich"

repeat_region

11992. .12028
/rpt_family="CAGA)n"

repeat_region

complement(13159. .15310)
/rpt_family="Lx6"

misc_feature

15028. .15033
/note="clone boundary
clone_end:SP6
site:ECORI"

repeat_region

15675. .15912
/rpt_family="CATATA)n"

repeat_region

complement(15915. .16226)
/rpt_family="MTC"

repeat_region

complement(16878. .16993)
/rpt_family="MTC"

repeat_region

17006. .18011
/rpt_family="ORR1B-int"

repeat_region

18012. .18383
/rpt_family="RUTR10"

repeat_region

18391. .18678
/rpt_family="ORR1B-int"

repeat_region

18675. .18730
/rpt_family="ORR1D"

repeat_region

18766. .18942
/rpt_family="CATATA)n"

repeat_region

18992. .19178
/rpt_family="CATATA)n"

repeat_region

19227. .19408
/rpt_family="RUTR1"

repeat_region

19426. .19743
/rpt_family="RUTR1"

repeat_region

complement(19801. .20141)
/rpt_family="MT2B"

repeat_region

20411. .20809
/rpt_family="RUTR9C"

repeat_region

complement(20817. .23539)
/rpt_family="MERVL"

repeat_region

complement(23544. .23788)
/rpt_family="MT2A"

repeat_region

23795. .24137
/rpt_family="ORR1C"

repeat_region

complement(24138. .24444)
/rpt_family="L1"

repeat_region

24445. .24486
/rpt_family="L1"

repeat_region

24895. .25932
/rpt_family="TG)n"

repeat_region

complement(24487. .24525)
/rpt_family="L1"

repeat_region

25978. .30421
/rpt_family="Lx2B"

repeat_region

complement(30422. .30569)
/rpt_family="B1_MM"

repeat_region

30570. .31592
/rpt_family="L1_MM"

repeat_region

complement(31593. .31918)

Query Match 76.8%; Score 19.2; DB 10; Length 219990;
Best Local Similarity 87.5%; Pred. No. 1e+02; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AGGCAGTAGTTGCCACACCTATGG 25
|||||

DB 198507 AGGCAGTAGTGGCCACACCTTAG 198484
|||||

Search completed: December 15, 2004, 00:42:58
Job time : 88.2037 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 21:00:04 : Search time 11.7084 Seconds

(without alignments)
11208.653 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25
Sequence: 1 gagcgagtagtgcacacactatg 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1980s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	6	ABT04897 Human G p
2	25	100.0	1202	6	ABV73364 Human TGR
3	25	100.0	1343	4	AAS15731 Aas15731 DNA encod
4	25	100.0	1343	10	ADJ87768 G-coupled
5	25	100.0	1343	12	ADT79324 NOV11 cod
6	25	100.0	1343	12	ADT79324 NOV11 cod
7	25	100.0	1526	6	AAD46858 Human 7TM
8	25	100.0	1526	8	ACA60889 Human CDN
9	25	100.0	1526	10	AAS57024 Aas57024 cDNA enco
10	25	100.0	1826	6	ABN84272 Human che
11	25	100.0	2273	12	ADO28955 Human nov
12	18.2	72.8	295	4	AAS7320 Human pan
13	18.2	72.8	147309	6	ABK49450 Human tta
14	17.6	70.4	879	4	AAL94220 Human neu
15	17.6	70.4	1376	12	ADO60492 ADO60492 Chlamydom
16	17.6	70.4	152501	12	ADP67269 Human chr
17	17.4	69.6	208	5	ADT73980 Human ova
18	17.4	69.6	208	5	ADT73980 Human ova
19	17.4	69.6	248	5	ADL45408 Human ova
20	17.4	69.6	350	3	AAC08203 Human sec
21	17.4	69.6	527	5	ADL39217 Human ova

C 22	17.4	69.6	674	4	AAS31458 Aas31458 Human CDN
C 23	17.4	69.6	674	6	ABO66782 Abo66782 Human pol
C 24	17.4	69.6	674	10	ADL10804 Adl10804 Human CDN
C 25	17.4	69.6	1182	6	ABZ22351 Abz22351 BM hepara
C 26	17.4	69.6	1596	10	ADP68972 Adp68972 Avian pne
C 27	17.4	69.6	1596	10	ADP68972 Adp68972 Avian pne
C 28	17.4	69.6	1596	12	ADJ96792 Adj96792 Avian pne
C 29	17.4	69.6	1596	12	ADJ96792 Adj96792 Avian pne
C 30	17.4	69.6	1596	12	ADM67263 Adm67263 Avian pne
C 31	17.4	69.6	1614	3	AAAI4952 Aaa14952 Nucleotid
C 32	17.4	69.6	1644	12	ADQ30909 Adq30909 Avian pne
C 33	17.4	69.6	1666	10	ADP68973 Adp68973 Avian pne
C 34	17.4	69.6	1666	10	ADP68973 Adp68973 Avian pne
C 35	17.4	69.6	1666	12	ADJ96793 Adj96793 DNA of th
C 36	17.4	69.6	1666	12	ADJ96793 Adj96793 DNA of th
C 37	17.4	69.6	1666	12	ADM67264 Adm67264 Avian pne
C 38	17.4	69.6	1860	10	ADP68975 Adp68975 Avian pne
C 39	17.4	69.6	1860	10	ADP68975 Adp68975 Avian pne
C 40	17.4	69.6	1860	12	ADJ96795 Adj96795 Avian pne
C 41	17.4	69.6	1860	12	ADJ96795 Adj96795 Avian pne
C 42	17.4	69.6	1860	12	ADM67266 Adm67266 Avian pne
C 43	17.4	69.6	2162	4	AAS31353 Aas31353 Human CDN
C 44	17.4	69.6	2162	6	ABO66677 Abo66677 Human pol
C 45	17.4	69.6	2162	10	ADC10699 Adc10699 Human CDN

ALIGNMENTS

RESULT 1	ABT04897	ABT04897 standard; DNA; 25 BP.
AC	ABT04897	
AC	ABT04897	
DT	11-OCT-2002 (first entry)	
DE	Human G protein coupled receptor hrup35 PCR primer SEQ ID NO: 42.	
KW	Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31; hrup32; hrup33; hrup34; hrup35; hrup36; hrup37; PCR; primer; ss.	
OS	Homo sapiens.	
XX	WO200242461.A2.	
XX	30-MAY-2002.	
XX	26-NOV-2001; 2001WO-US044386.	
XX	27-NOV-2000; 2000US-0253404P.	
XX	12-DEC-2000; 2000US-0255366P.	
XX	20-FEB-2001; 2001US-0270266P.	
XX	20-FEB-2001; 2001US-0270266P.	
XX	06-APR-2001; 2001US-0282032P.	
XX	06-APR-2001; 2001US-0282356P.	
XX	06-APR-2001; 2001US-0282358P.	
XX	06-APR-2001; 2001US-0282358P.	
XX	14-MAY-2001; 2001US-0290917P.	
XX	31-JUL-2001; 2001US-0309208P.	
XX	(AREN-) ARENA PHARM INC.	
XX	Chen R, Chu ZL, Dang HT, Lowitz KP, Fride C;	
XX	WPI; 2002-566565/60.	
XX	Novel endogenous and non-endogenous versions of G protein-coupled	
XX	receptor useful for identification of candidate compounds as receptor	
XX	agonists or antagonists for use as therapeutic agents.	
XX	Example 1; Page 28; 84pp; English.	

CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a PCR primer used to isolate a GPCR coding sequence of the invention
XX
XX Sequence 25 BP; 6 A; 6 C; 8 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.024; 0; Indels 0; Gaps 0;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGGCCACACCTATGG 25
Db 1 GAGGCACTAGTGGCCACACCTATGG 25

RESULT 2
ABV73364/C
ID ABV73364 standard; DNA; 1202 BP.

XX
XX ABV73364;

XX 22-JAN-2003 (first entry)

XX Human TGR20 polypeptide encoding DNA.

XX G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;
XX antileptic; nootropic; neuroprotective; antiemetic; antitumor; human;
XX antiparkinsonian; antileptic; TGR20; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 68..1129
XX FT /*tag= a
XX FT /product= "TGR20"
XX FT /note= "GPCR polypeptide"

XX WO200277001-A2.

XX 03-OCT-2002.

XX 08-MAR-2002; 2002WO-US007171.

XX 09-MAR-2001; 2001US-00802803.

XX 16-MAR-2001; 2001US-0276649P.

XX (TULAR) TULARIK INC.

XX Tian H, Zhao J, Chen J, Cutler G;

XX WPI; 2003-018681/01.

XX P-PSDB; ABB82499.

XX New G-protein coupled receptor polypeptides and polynucleotides useful
XX for identifying compounds for treating a TGR-associated disorder, e.g.
XX psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's
XX disease, anemia.

XX Claim 6; Page 63; 87pp; English.

XX The invention relates to G-protein coupled receptor (GPCR) polypeptides
XX and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,
XX TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are
XX useful for identifying compounds for treating a TGR-associated disorder,
XX such as psoriasis, inflammatory bowel disease, hyperlipidemia,
XX Parkinson's disease, Huntington's disease, anemia, immune and blood
XX disorders, ulcerative colitis, Crohn's disease or spleen enlargement.
XX They are also useful for identifying cells such as kidney, liver,
XX hypothalamus, colon, adipose, or spleen cells, for forensics and
XX paternity determination, diagnosing diseases and examining signal
XX transduction. The present sequence represents a human TGR20 polypeptide

CC encoding DNA
XX
XX Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 8; Length 1202;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGGCCACACCTATGG 25
Db 1156 GAGGCACTAGTGGCCACACCTATGG 1132

RESULT 3
AAS15731/C
ID AAS15731 standard; CDNA; 1343 BP.

XX
XX AAS15731;

XX 14-FEB-2002 (first entry)

XX DNA encoding chemokine receptor family related protein, NOV11.

XX NOV; cytoskeletal; psoriasis; nootropic; neuroprotectant;
XX cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;
XX haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;
XX differentiation; proliferation; haematopoiesis; wound healing;
XX angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;
XX haemophilia; allergy; pendred syndrome; skeletal dysplasia;
XX ischaemic injury; neuroepithelial disorder; hepatitis; heart failure;
XX chemokine receptor; chromosome 1; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX 5'UTR 1
XX FT /*tag= a
XX FT 2..1063
XX FT /*tag= b
XX FT /product= "Chemokine receptor related protein, NOV11"

XX sig_peptide 2..142 C
XX FT /*tag= 143..1060
XX FT mat_peptide 143..1060
XX FT /*tag= d
XX FT /note= "Mature chemokine receptor related protein, NOV11"

XX 3'UTR 1064..1343
XX FT /*tag= e

XX WO200170978-A2.

XX 27-SEP-2001.

XX 20-MAR-2001; 2001WO-US009093.

XX 20-MAR-2000; 2000US-0190768P.

XX 20-MAR-2000; 2000US-0190835P.

XX 22-MAR-2000; 2000US-0190972P.

XX 22-MAR-2000; 2000US-0191199P.

XX 24-MAR-2000; 2000US-0191947P.

XX 28-MAR-2000; 2000US-0192657P.

XX 28-MAR-2000; 2000US-0192664P.

XX 28-MAR-2000; 2000US-0192665P.

XX 28-MAR-2000; 2000US-0192844P.

XX 29-MAR-2000; 2000US-0192836P.

XX 31-MAR-2000; 2000US-0193843P.

XX (CURA-) CURAGEN CORP.

XX Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;

XX Vernet CM;

XX WPI; 2001-639127/73.

XX P-PSDB; AAU10066.

XX polypeptides and nucleic acids related to chloride channel, insulin-like
PT growth factor family of proteins, useful for diagnosing and treating
PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.
XX
XX Claim 9; Page 44; 151pp; English.
XX
CC The invention describes isolated NOVX (NOVX1-11) polypeptides, NOVX
CC polypeptides are useful for treating pathology associated with NOVX
CC polypeptide, determining the presence of or predisposition to a disease
CC associated with altered levels of NOVX, identifying agents binding to
CC NOVX and treatment of disorders associated with altered expression of
CC members of chloride channel-associated proteins e.g. cystic fibrosis and
CC congenital myotonia. NOVX proteins are useful in treatment of disorders
CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,
CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells
CC in one or more organs (e.g. haemophilia, anaemia), Pendred syndrome,
CC skeletal dysplasias, disorders characterized by altered cell shape,
CC motility, and apoptosis, ischemic injury, hepatitis, neuroepithelial
CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the
CC treatment of disorders of vascular smooth muscle cell differentiation,
CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are
CC useful to screen for molecules which inhibit or enhance NOVX activity or
CC function and are useful as targets for the identifying small molecules,
CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell
CC differentiation, haemostasis, wound healing and angiogenesis. NOV
CC sequences are also useful for: identifying a cell or tissue type in a
CC biological sample; amplifying DNA sequences from very small biological
CC samples e.g. hair or skin or body fluids and as primers and probes to
CC identify and/or clone NOVX homologues. NOVX proteins are useful
CC immunogens to generate antibodies to monitor protein levels and modulate
CC NOVX activity. Cells comprising the nucleic acids are useful for
CC producing transgenic animals, for studying the function and/or activity
CC of NOVX protein and identifying and/or evaluating modulators of NOVX
CC protein activity. This sequence encodes NOV11 (located on chromosome 1)
CC related to the chemokine receptor family of proteins, one of 12 NOV
CC polypeptides described in the method of the invention
XX
SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 25; DB 4; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GAGGCACTAGTGGCCACACTATGG 25
Db 1090 GAGGCACTAGTGGCCACACTATGG 1066
XX
RESULT 4
ADJ87768/c
ID ADJ87768 standard; DNA; 1343 BP.
XX
AC ADJ87768;
XX
DT 06-MAY-2004 (first entry)
XX
DE G-coupled protein receptor-related protein coding sequence #57.
XX
XX novel protein; G-coupled protein receptor-related protein;
KM cardiomyopathy; atherosclerosis; cell signal processing-related disorder;
KM metabolic pathway modulation-related disorder; diabetes; cancer; stroke;
KM Huntington's disease; epilepsy; pain; hypercholesterolaemia;
KM obesity; hypertension; Crohn's disease; systemic lupus erythematosus;
KM viral infections; bacterial infection; parasitic infection;
KM hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;
KM Alzheimer's disease; tuberculous sclerosis; hypercalcaemia; cerebral palsy;
KM gene; ds.
XX
XX Undeidentified.
XX
XX OS
XX PN
XX MO2002102321-A2.

PD 27-DEC-2002.
XX
XX 18-JUN-2002; 2002MO-US019522.
PF
XX
XX 18-JUN-2001; 2001US-0298994P.
PR 18-JUN-2001; 2001US-0299134P.
PR 04-OCT-2001; 2001US-00972446.
PR 06-JUN-2002; 2002US-00299134.
PR 07-JUN-2002; 2002US-00298994.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;
PI Pena CE, Spaderna SK, Zhong M;
XX WPI; 2003-167441/16.
DR P-PSDB; ADJ87769.
XX
XX New MOLX polypeptides and polynucleotides, useful in gene therapy,
PT particularly for treating or preventing e.g. cardiomyopathy,
PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate
PT cancer, stroke or pain.
XX
XX Claim 8; SEQ ID NO 203; 378pp; English.
XX
CC The invention comprises the amino acid and coding sequences of novel G-
CC coupled protein receptor-related (MOL) proteins. The DNA and protein
CC sequences of the invention are useful for treating or preventing a MOL-
CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders
CC associated with cell signal processing and metabolic pathway modulation,
CC or diabetes. The DNA and protein sequences are also useful for the
CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,
CC pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,
CC systemic lupus erythematosus, viral infections, bacterial infections,
CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau
CC syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or
CC cerebral palsy. The present DNA sequence encodes a MOL protein of the
CC invention.
XX
SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 25; DB 10; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GAGGCACTAGTGGCCACACTATGG 25
Db 1090 GAGGCACTAGTGGCCACACTATGG 1066
XX
RESULT 5
ADJ79324/c
ID ADJ79324 standard; DNA; 1343 BP.
XX
XX ADJ79324;
AC
XX
DT 22-APR-2004 (first entry)
XX
DE NOV11 coding sequence, SEQ ID 21.
XX
XX Cytotoxic; Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;
KM Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
KM Ophthalmological; Antiproliferative; Neuroprotective; Nootropic;
KM Antiatherosclerotic; Hypotensive; Cardiant; Cerebroprotective;
KM Gene Therapy; NOVX; human; cancer; myelogenous leukemia;
KM congenital neonatal autoimmune thrombocytopenia; immunological disorder;
KM allergy; infection; asthma; lung disease; reproductive disorder;
KM haemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;
KM diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;
KM hypertension; stroke; heart failure; chromosome 1; NOV11;
KM chemokine receptor; gene; ds.
XX
XX OS
XX Homo sapiens.

```

XX Key Location/Qualifiers
FH 2.1063
FT CDS /tag= a
FT /product= "NOV11"
XX
XX WO2004009635-A2.
XX
XX 29-JAN-2004.
XX
XX 04-OCT-2001; 2001WO-US031292.
XX
XX 20-MAR-2001; 2001US-00813432.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Tanipler RJ, Majumder K, Spaderma SK, Smithson G, Mezes PS,
XX Vernet CAM;
XX
XX WPI; 2004-123380/12.
XX
XX P-PSDB; ADI79325.
XX
XX Claim 8; Page 43-44; 158pp; English.
XX
XX The present invention relates to novel NOVX proteins and their coding
XX sequences (ADI79304-ADI79327). The sequences are useful for the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease associated with the protein, or for diagnosing and treating
XX disorders associated with the NOVX protein, such as cancer, myelogenous
XX leukaemia, congenital neonatal autoimmune thrombocytopenia,
XX immunological disorders, allergy and infection, asthma, lung diseases,
XX reproductive disorders, male and female reproductive diseases,
XX haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,
XX diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
XX hypertension, stroke and heart failure. NOV11 represents a new member of
XX the chemokine receptor family and the gene is located on chromosome 1.
XX
XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 25; DB 12; Length 1343;
XX Best Local Similarity 100.0%; Pred. No. 0.048;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GAGGCACTAGTGGCCACACCTATGG 25
XX |||||||||||||||||||||||
XX 1090 GAGGCACTAGTGGCCACACCTATGG 1066
XX
XX RESULT 6
XX ADO56003/c
XX ID ADO56003 standard; cDNA; 1343 BP.
XX
XX AC ADO56003;
XX
XX 15-JUL-2004 (first entry)
XX
XX DNA encoding human NOV11.
XX
XX human; gene; ss; cancer; obesity; diabetic nephropathy;
XX acute pancreatitis; stroke; multiple sclerosis.
XX
XX Homo sapiens.
XX
XX US2004058862-A1.
XX
XX 25-MAR-2004.
XX
XX 18-SEP-2002; 2002US-00246583.
XX
XX 18-SEP-2002; 2002US-00246583.
XX
XX (MAJU/) MAJUMDER K.
XX

```

```

PI Majumder K;
XX
XX WPI; 2004-268835/25.
XX
XX P-PSDB; ADO56004.
XX
XX Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic
XX nephropathy, acute pancreatitis, strokes and multiple sclerosis.
XX
XX Disclosure; Page 28; 87pp; English.
XX
XX The invention relates to novel isolated NOVX nucleic acids and encoded
XX polypeptides. The nucleic acids, polypeptides and antibodies raised
XX against the polypeptides are useful for preventing or treating diseases
XX associated with aberrant NOVX expression or activity e.g., cancer,
XX obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple
XX sclerosis. The present sequence represents a NOVX nucleic acid of the
XX invention.
XX
XX Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 25; DB 12; Length 1343;
XX Best Local Similarity 100.0%; Pred. No. 0.048;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GAGGCACTAGTGGCCACACCTATGG 25
XX |||||||||||||||||||||||
XX 1090 GAGGCACTAGTGGCCACACCTATGG 1066
XX
XX RESULT 7
XX AAD46858/c
XX ID AAD46858 standard; cDNA; 1526 BP.
XX
XX AC AAD46858;
XX
XX 27-JAN-2003 (first entry)
XX
XX Human 7TM domain receptor 65507 cDNA.
XX
XX Human; adenosine deaminase; seven transmembrane domain receptor; cancer;
XX 7TM; glycoprotease; immune disorder; IGA deficiency; allergy; arhythmia;
XX rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;
XX hypertension; ischaemic heart disease; obesity; myocardial infarction;
XX endothelial cell disorder; Grave's disease; psoriasis; brain disorder;
XX Parkinson's disease; Alzheimer's disease; haematopoietic disorder;
XX cerebral oedema; metabolic disorder; liver disorder; platelet disorder;
XX chromosome mapping; tissue typing; gene therapy; neuroprotective;
XX cytostatic; anorectic; cardiant; haemostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX Key Location/Qualifiers
XX CDS 139..1200
XX
XX /tag= a
XX /product= "Human 65507 protein"
XX /note= "This region is specifically claimed as SEQ ID NO:
XX 12 in claim 1 of the specification"
XX
XX sig_peptide 139..306
XX
XX mat_peptide 307..1197
XX
XX /tag= b
XX /tag= c
XX /product= "Human mature 65507 protein"
XX
XX WO200274960-A2.
XX
XX 26-SEP-2002.
XX
XX 08-NOV-2001; 2001WO-US051427.
XX
XX 08-NOV-2000; 2000US-0246768P.
XX
XX 08-NOV-2000; 2000US-0246772P.
XX
XX 15-NOV-2000; 2000US-0249185P.
XX

```

PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Leiby KR, Kapeller-Libermann R, Glucksmann M;
 XX
 DR WPI; 2002-755989/82.
 DR P-PSDB; AAE29236.
 XX
 PT New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,
 PT useful for diagnosing and treating cancer, immune, cardiovascular,
 PT hematopoietic, brain, pain, metabolic, liver or platelet disorders, and
 PT in pharmacogenomics.
 XX
 PS Claim 1; Fig 17; 178pp; English.
 XX
 CC The present invention relates to novel 38650, 28472, 5495, 65507, 81588
 CC or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-
 CC protease or seven transmembrane domain (7TM) receptor family members.
 CC Sequences of the invention are useful in diagnosing and treating cancer
 CC or aberrant cellular proliferation and/or differentiation (e.g. colon or
 CC lung cancer), immune disorders (e.g. selective IgA deficiency, rheumatoid
 CC arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g.,
 CC hyperension, atherosclerosis, arrhythmias, ischaemic heart disease,
 CC myocardial infarction, thrombus) including endothelial cell disorders
 CC (e.g. psoriasis, Grave's disease), hematopoietic disorders, brain
 CC disorders (e.g. cerebral edema, Parkinson's or Alzheimer's disease),
 CC pain and metabolic disorders (e.g. obesity), liver disorders or platelet
 CC disorders. They are also useful in screening assays, predictive medicine
 CC (e.g. diagnostic assays, prognostic assays, monitoring clinical trials
 CC and pharmacogenetics) and prophylactic and therapeutic methods. The
 CC nucleic acids may also be used in chromosome mapping, tissue typing and
 CC forensic biology and as surrogate markers. Sequences of the invention are
 CC also used in gene therapy. The present sequence is human 7TM domain
 CC receptor 65507 CDNA
 XX
 SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 Query Match 100.0%; Score 25; DB 6; Length 1526;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGGCACTAGTGGCACACCTATGG 25
 Db 1227 GAGGCACTAGTGGCACACCTATGG 1203
 RESULT 8
 ID ACA60889/c
 AC ACA60889 standard; cDNA; 1526 BP.
 XX
 AC ACA60889;
 XX
 DT 08-JUL-2003 (first entry)
 XX
 DE Human cDNA encoding 65507 protein.
 XX
 KW Human; ser; gene; cancer; aberrant cellular proliferation;
 KW differentiation; immune disorders; heart disorder; brain disorder;
 KW cardiovascular disorder; endothelial cell disorder; pain disorder;
 KW hematopoietic disorder; blood vessel disorder; metabolic disorder;
 KW liver disorder; platelet disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 139..1200
 FT /*tag= a
 FT /product= "Protein 65507"
 FT /note= "This CDS is specifically claimed in claim 1"
 XX
 PN US2003009017-A1.
 XX
 PD 09-JAN-2003.
 XX

PF 08-NOV-2001; 2001US-00012140.
 XX
 PR 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX
 PA (LEIB/) LEIBY K R.
 PA (KAP/) KAPPELLER-LIBERMANN R.
 PA (GLUC/) GLUCKSMANN M A.
 XX
 PI Leiby KR, Kapeller-Libermann R, Glucksmann MA;
 XX
 DR WPI; 2003-42888/40.
 DR P-PSDB; ABU09571.
 XX
 PT New isolated 38650, 28472, 5495, 65507, 81588 and 14354 nucleic acid
 PT molecules, useful for diagnosing, treating cancer, pain, or immune,
 PT heart, endothelial cell, hematopoietic, blood vessel, brain, metabolic
 PT and liver disorders.
 XX
 PS Claim 2; Fig 17; 90pp; English.
 XX
 CC The invention relates to an isolated 38650 (encoding adenosine
 CC deaminase), 28472 (encoding a glycoprotease), 5495 (encoding a 7
 CC transmembrane receptor), 65507, 81588 or 14354 nucleic acid molecule (or
 CC a sequence which is at least 60% identical to the six nucleic acids or
 CC their open reading frames, fragments of at least 15 nucleotides,
 CC naturally occurring variants, or a DNA insert of the plasmid deposited
 CC with the American Type Culture Collection as Accession No. not defined in
 CC the specification, which encodes the amino acid sequence). Also included
 CC are a host cell containing the nucleic acids (used to produce the
 CC proteins), the encoded proteins, an antibody that selectively binds to
 CC the polypeptide, and identifying a compound that binds to/modulates the
 CC activity of the polypeptide. The nucleic acid molecules, polypeptides and
 CC methods are useful for diagnosing, treating cancer, aberrant cellular
 CC proliferation and/or differentiation, immune disorders, heart disorders,
 CC cardiovascular disorders including endothelial cell disorders,
 CC hematopoietic disorders, blood vessel disorders, brain disorders, pain
 CC and metabolic disorders, liver disorders and platelet disorders (many
 CC examples of these disorders are given in the specification). The present
 CC sequence is the Human CDNA 65507
 XX
 SQ Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 Query Match 100.0%; Score 25; DB 8; Length 1526;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGGCACTAGTGGCACACCTATGG 25
 Db 1227 GAGGCACTAGTGGCACACCTATGG 1203
 RESULT 9
 ID ABS57024/c
 AB ABS57024 standard; cDNA; 1526 BP.
 XX
 AC ABS57024;
 XX
 DT 30-JAN-2003 (first entry)
 XX
 DE cDNA encoding novel human 7 transmembrane receptor 65507.
 XX
 KW Cancer; aberrant cell proliferation; aberrant cell differentiation;
 KW breast cancer; ovarian cancer; prostate cancer; colon cancer;
 KW lung cancer; immune disorder; heart disorder; cardiovascular disorder;
 KW endothelial disorder; hematopoietic disorder; blood vessel disorder;
 KW brain disorder; pain; metabolic disorder; liver disorder; diabetes;
 KW platelet disorder; carcinoma; sarcoma; leukaemia; Hodgkin's disease;
 KW autoimmune disorder; hypertension; atherosclerosis; heart failure;
 KW myocardial infarction; ischaemic heart disease; Crohn's disease;
 KW Grave's disease; Kawasaki syndrome; Raynaud's disease; aneurysm;
 KW cerebral ischaemia; peripheral neuropathy; Alzheimer's disease;

KW Parkinson's disease; anorexia nervosa; cachexia; gene; ss;
 XX 7 transmembrane receptor.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 139..1200
 FT /tag= a
 FT /product= "7 transmembrane receptor 65507"
 FT /note= "Specifically claimed in claim 1"
 PN WO200277233-A2.
 XX
 XX
 PD 03-OCT-2002.
 PF 08-NOV-2001; 2001WO-US046724.
 XX
 XX
 PR 08-NOV-2000; 2000US-0246768P.
 PR 08-NOV-2000; 2000US-0246772P.
 PR 15-NOV-2000; 2000US-0249185P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI LeiJy KR, Kapeller-Libermann R, Glucksmann M;
 DR WPI; 2003-029938/02.
 DR P-PSDB; ABB71163.
 XX
 PT New adenosine deaminase, glycoprotease and seven transmembrane domain
 PT nucleic acids and polypeptides, designated 38650, 28472, 5495, 65507,
 PT 81588 and 14354, useful for treating e.g. leukemias, Hodgkin's disease or
 PT hypertension.
 XX
 PS Claim 2; Fig 17; 178pp; English.
 XX
 XX The invention describes isolated 38650, 28472, 5495, 65507, 81588 and
 CC 14354 nucleic acid molecules (I) and their encoded polypeptides (II). The
 CC 38650 nucleic acid molecule comprises a sequence encoding adenosine
 CC deaminase, the 28472 nucleic acid molecule comprises a sequence encoding
 CC a human glycoprotease. The 5495, 65507, 81588 and 14354 comprise
 CC sequences that encode a human seven transmembrane domain (7TM). The
 CC 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid and polypeptide
 CC sequences are useful for diagnosing, preventing or treating a subject
 CC with or at risk of developing a disorder, e.g. cancer or aberrant
 CC cellular proliferation and/or differentiation (e.g. breast, ovarian,
 CC prostate, colon or lung cancer), immune disorders, heart disorders,
 CC cardiovascular disorders, endothelial disorders, hematopoietic disorders,
 CC blood vessel disorders, brain disorders, pain and metabolic disorders,
 CC liver disorders or platelet disorders. These disorders include carcinoma,
 CC sarcoma, leukaemia, Hodgkin's disease, autoimmune disorders,
 CC hypertension, atherosclerosis, heart failure, myocardial infarction,
 CC ischaemic heart disease, Crohn's disease, Grave's disease, Kawasaki
 CC syndrome, Reynaud's disease, aneurysm, cerebral ischaemia, peripheral
 CC neuropathy, Alzheimer's disease, Parkinson's disease, anorexia nervosa,
 CC cachexia or diabetes. This sequence encodes the novel human 7
 CC transmembrane receptor 65507
 XX
 SO Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;
 Query Match 100.0%; Score 25; DB 10; Length 1526;
 Best Local Similarity 100.0%; Pred. No. 0.049; Indels 0; Gaps 0;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGGCGTAGTTCGCCACCTATGG 25
 Db 1227 GAGGCGTAGTTCGCCACCTATGG 1203
 RESULT 10
 ABN84272/c
 ID ABN84272 standard; cDNA, 1826 BP.
 XX
 AC ABN84272;

XX 23-SEP-2002 (first entry)
 DT
 XX
 XX Human chemokine-like receptor cDNA.
 DE
 XX
 XX Chemokine-like receptor; G-protein coupled receptor; receptor; human;
 KW HIV infection; cardiovascular disease; asthma;
 KW chronic obstructive pulmonary disease; cardiac; antiasthmatic;
 KW vasotropic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;
 KW antiinflammatory; antiallergic; immunomodulator; gene therapy; gene; ss.
 XX
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 36..1105
 FT /tag= a
 FT /product= "Chemokine-like receptor"
 FT /transl_except= (pos:51..52;aa:Leu)
 FT /note= "apparent 1 base frameshift"
 XX
 XX WO200248358-A2.
 PN
 XX
 XX
 PD 20-JUN-2002.
 PF 12-DEC-2001; 2001WO-BP014571.
 XX
 XX
 PR 14-DEC-2000; 2000US-0255150P.
 PR 02-APR-2001; 2001US-0280110P.
 PR 21-JUN-2001; 2001US-0299474P.
 XX
 XX (FARB) BAYER AG.
 PA
 XX
 PI Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;
 XX
 XX WPI; 2002-547858/58.
 DR P-PSDB; ABB79518.
 XX
 PT New isolated polynucleotide encoding a chemokine-like receptor
 PT polypeptide for treating e.g. asthma, myocardial infarction, human
 PT immunodeficiency virus infection, or chronic obstructive pulmonary
 PT disease.
 XX
 PS Disclosure; Fig 6; 114pp; English.
 XX
 XX The present sequence is that of cDNA encoding a novel human chemokine-
 CC like receptor (see AMN79518) of 356 amino acids (41.4 kDa) and having 7
 CC putative transmembrane domains, consistent with the structure of a G-
 CC protein coupled receptor. Its closest human homologue is C-C chemokine
 CC receptor 3. The novel receptor is expressed at low levels in most
 CC tissues. It is expressed at a high level in phytohaemagglutinin-
 CC stimulated CD8+ cells, but in none of the other immune cells tested. It
 CC may act as a receptor of chemottractant molecules on activated
 CC lymphocytes and be involved in cell trafficking and homing to sites of
 CC infection, inflammation or tissue injury. Regulation of activity of the
 CC novel receptor can therefore be used to treat cardiovascular,
 CC immunological and inflammatory diseases, including asthma and chronic
 CC obstructive pulmonary disease (COPD). The receptor may also be a target
 CC for viruses that reside in the nervous system. Regulating the binding of
 CC ligands, e.g. chemottractant molecules or virus particles, to the
 CC receptor can therefore be used to modulate the immune response to inhibit
 CC viral infections, including HIV infection. A claimed method of reducing
 CC activity of the receptor involves contacting a cell with a reagent
 CC (preferably an antibody, antisense oligonucleotide or ribozyme) to a
 CC product (preferably RNA or a polypeptide) encoded by a polynucleotide
 CC encoding the human chemokine-like receptor in vivo or in vitro. A claimed
 CC method of treating a chemokine-like receptor dysfunction related disease
 CC selected from HIV infection, a cardiovascular disorder, asthma or COPD
 CC uses a reagent that modulates a function of the receptor
 XX
 SO Sequence 1826 BP; 452 A; 510 C; 371 G; 493 T; 0 U; 0 Other;
 Query Match 100.0%; Score 25; DB 6; Length 1826;
 Best Local Similarity 100.0%; Pred. No. 0.05;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGGCACCTATGG 25
 |||||
 Db 1132 GAGGCACTAGTGGCACCTATGG 1108

RESULT 11
 ADO28955/C
 ID ADO28955 standard; cDNA; 2273 BP.

AC ADO28955;
 XX
 DT 29-JUL-2004 (first entry)

DE Human novel GPCR PCR3 polynucleotide, SEQ ID NO:54.
 XX

KM G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KM transgenic mouse; neurological disorder; adrenal gland disorder;
 KM colon disorder; intestinal disorder; cardiovascular disorder;
 KM muscular disorder; blood disorder; immune disorder; bone disorder;
 KM joint disorder; metabolic disorder; nutritive disorder; cancer;
 KM kidney disorder; liver disorder; lung disorder; breast disorder;
 KM ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KM skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KM thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
 KM cytoactive; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
 KM CNS; central nervous system; respiratory; antidiabetic; antidiabetic;
 KM virologic; hepatotropic; antibacterial; antianemic; antiseborrheic;
 KM dermatological; antitumor; antihypertensive; anorectic;
 KM immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
 KM gene; ss.

OS Homo sapiens.
 XX
 PN MO200404000-A2.
 XX
 PD 13-MAY-2004.

PF 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.

XX (PRIM-) PRIMAL INC.
 PA
 PI Galenaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Medisen L, McIlwain KL, Pavlova KM, Vasilevici D, Zeng H;
 PI WPI; 2004-350329/36.
 DR P-PSDB; ADO28954.
 DR
 XX

PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 PT

PS Claim 13; SEQ ID NO 54; 542P; English.
 XX
 XX

CC The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of

CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancer). The present sequence represents a GPCR-encoding
 CC nucleic acid of the invention. Note: The full sequence data for this
 CC patent did not form part of the printed specification; those sequences
 CC not shown were obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX

SEQ Sequence 2273 BP; 503 A; 691 C; 549 G; 530 T; 0 U; 0 Other;

Query Match 100.0%; Score 25; DB 12; Length 2273;
 Best Local Similarity 100.0%; Pred. No. 0.052;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTGGCACCTATGG 25
 |||||
 Db 1541 GAGGCACTAGTGGCACCTATGG 1517

RESULT 12
 AAH57320/C
 ID AAH57320 standard; cDNA; 295 BP.
 XX
 AC AAH57320;
 XX
 DT 10-SEP-2001 (first entry)
 XX

DE Human pancreas specific cDNA sequence SEQ ID NO:160.
 XX

KM Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
 KM liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KM metabolic disease; developmental disease; cystostatic; immunomodulatory;
 KM neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

XX Homo sapiens.
 OS
 XX
 PN MO200132927-A2.
 XX
 PD 10-MAY-2001.
 XX

PF 02-NOV-2000; 2000WO-US030396.
 XX
 PR 04-NOV-1999; 99US-0163508P.
 PR

XX (INCY-) INCYTE GENOMICS INC.
 PA
 PI Sornasse T, Sellhammer JJ, Watson GA;
 PI WPI; 2001-291057/30.
 DR
 XX

PT New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is
 PT associated with a cancer, immunopathology or neuropathology.
 PT

PS Claim 1; Page 125; 327P; English.
 XX
 XX

CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
 CC sequences (I). (I) can have cytosstatic, immunomodulatory and
 CC neuroprotective activities, and can be used in gene therapy. (I) and
 CC proteins (II) encoded by them are used in high throughput screening
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or their
 CC fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical

CC agents. Expression of (i) in a sample indicates the differentiation of
CC embryonic stem cells into a tissue selected from brain, heart, kidney,
CC liver, lung, skeletal muscle or pancreatic tissues. (i) and (ii) are used
CC to produce an expression profile that defines a metabolic or
CC developmental process, treatment, condition, disease or disorder. The
CC gene profile can be used for diagnosis, prognosis or monitoring of
CC treatments and for investigating a predisposition to a disorder where the
CC gene is associated with a cancer, immunopathology or neuropathology

Sequence 295 BP; 59 A; 82 C; 84 G; 60 T; 0 U; 10 Other;

Query Match 72.8%; Score 18.2; DB 4; Length 295;
Best Local Similarity 80.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 GAGGAGTGTGTCACACCTATGG 25
52 GANNACAGATGTGCCATCTCAATGG 28

RESULT 13
ABK49450
ID ABK49450 standard; DNA; 147309 BP.

XX ABK49450;

DT 15-JUL-2002 (first entry)

DE Human transporter genomic DNA sequence.

KW Human; transporter protein; therapeutic target; query sequence;
KW database search; single nucleotide polymorphism; SNP; chromosome 5; gene;
ds.

XX Homo sapiens.

XX Key Location/Qualifiers
FH exon 2002..2778

FT /tag= a
FT /number= 1
FT intron 2779..43601

FT /tag= b
FT /number= 1
FT variation replace(3356,T)

FT /tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(3396,C)

FT /tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(3493,A)

FT /tag= x
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(3780,GC)

FT /tag= y
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4063,G)

FT /tag= z
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4628,G)

FT /tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4702,A)

FT /tag= ab
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5235,C)

FT /tag= ac
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(5638,A)

FT /tag= ad
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(6248,C)

FT /tag= ae
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(21708,G)

FT variation replace(7845,A)
FT /tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(8690,G)
FT /tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(10023,A)
FT /tag= ah
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(10049,T)
FT /tag= ai
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(10300,A)
FT /tag= aj
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(10842,C)
FT /tag= ak
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(12545,G)
FT /tag= al
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(13192,T)
FT /tag= am
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14227,C)
FT /tag= an
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14304,G)
FT /tag= ao
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14535,C)
FT /tag= ap
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14580,A)
FT /tag= aq
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14594,G)
FT /tag= ar
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(14836,C)
FT /tag= as
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(15578,A)
FT /tag= at
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16079,G)
FT /tag= au
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16118,G)
FT /tag= av
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16643,A)
FT /tag= aw
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(16712,C)
FT /tag= ax
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17201,C)
FT /tag= ay
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17228,T)
FT /tag= az
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17893..17895,AG)
FT /tag= ba
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17895..17897,GG)
FT /tag= bb
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(15102,A)
FT /tag= bc
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(21708,G)

```

FT      /*tag= bd
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(22641,A)
FT      /*tag= de
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(24758,G)
FT      /*tag= bf
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(25059,C)
FT      /*tag= dg
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26084, .26086, CAAC)
FT      /*tag= bh
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(26318,G)
FT      /*tag= bi
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27203, .27205, CTTG)
FT      /*tag= bj
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27291, .27293, TKG)
FT      /*tag= bk
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(27403, .27405, AYT)
FT      /*tag= bl
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28025,C)
FT      /*tag= bm
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28389,G)
FT      /*tag= bn
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(28587,G)
FT      /*tag= bo
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(29581,A)
FT      /*tag= bp
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(29727,G)
FT      /*tag= bq
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(30314,A)
FT      /*tag= br
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(30888,G)
FT      /*tag= bs
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37895,G)
FT      /*tag= bt
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37150,A)
FT      /*tag= bu
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37222,T)
FT      /*tag= bv
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(36395, .36397, ATTA)
FT      /*tag= bw
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37006,A)
FT      /*tag= bx
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(37919,A)
FT      /*tag= by
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(38891,T)
FT      /*tag= bz
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39293,G)
FT      /*tag= ca
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39933,G)
FT      /*tag= cb

```

```

FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(39968,A)
FT      /*tag= cc
FT      /*tag= cc
Query Match
Best Local Similarity 72.8%; Score 18.2; DB 6; Length 147309;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGCAGTAGTTCGCACCTAT 23
Db      122948 GAGCAGTAGTTCGCACCTAT 122970

RESULT 14
AA194220/C
ID      AA194220 standard; cDNA; 879 BP.
XX
XX
AC      AA194220;
XX
DT      13-NOV-2001 (first entry)
XX
DE      Human neuroblastoma expressed polynucleotide SEQ ID NO 295.
XX
KM      Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS      Homo sapiens.
XX
XX      WO200166719-A1.
XX
PD      13-SEP-2001.
XX
PF      02-MAR-2001; 2001WO-JP001629.
XX
PR      07-MAR-2000; 2000JP-00159195.
XX
PA      (CHIB-) CHIBA PHARMACEUTICAL.
XX      (HISM) HISAMITSU PHARM CO LTD.
XX
PI      Nakagawara A;
XX
XX      WPI; 2001-565584/63.
XX
XX      Nucleic acids originating in gene expressed in human neuroblastoma,
PT      useful as probe or primer in diagnosing prognosis of human neuroblastoma,
PT      malignancy and susceptibility indicator or tumor marker for anti-cancer
PT      agents.
XX
XX      Claim 1; Page 258; 2979pp; Japanese.
XX
XX      The invention relates to novel genes (AA193926-AA197963) expressed in
CC      human neuroblastoma. The nucleic acids are applicable as a probe or
CC      primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC      susceptibility indicators or tumor markers for anti-cancer agents. The
CC      gene information for diagnosing prognosis is related to factors similar
CC      to that for N-myc and TrkA genes
XX
SQ      Sequence 879 BP; 200 A; 226 C; 208 G; 216 T; 0 U; 29 Other;

Query Match
Best Local Similarity 70.4%; Score 17.6; DB 4; Length 879;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 AGCAGTAGTTCGCACCTATG 25
Db      441 AGCAGTAGTTCGCACCTATG 418

RESULT 15
ADO60492
ID      ADO60492 standard; DNA; 1376 BP.
XX
XX      ADO60492;
AC
XX

```

DT 15-JUL-2004 (first entry)
 XX Chlamydomonas ascorbate peroxidase gene SegId1.
 DE
 XX plant; grass family; cold-resistance; active oxygen elimination;
 KM ascorbate peroxidase; Chlamydomonas; chloroplast transfer sequence;
 KM translocated; chloroplast;
 KM thylakoid membrane proximity transfer sequence; food crop;
 KM cold region cultivation; gene; ds.
 XX
 OS Chlamydomonas.
 XX
 FH Key Location/Qualifiers
 FT CDS 32..991
 FT /*tag= a
 FT /product= "Chlamydomonas ascorbate peroxidase"
 XX
 XX JP2004105136-A.
 XX
 XX 08-APR-2004.
 XX
 XX 20-SEP-2002; 2002JP-00274942.
 XX
 XX 20-SEP-2002; 2002JP-00274942.
 XX
 XX (SHOK-) SHOKUBUTSU KOGAKU KK.
 XX (KANT) KANSAI DENRYOKU KK.
 XX
 XX WPI: 2004-289542/27.
 XX P-PSDB; ADO60493.
 XX
 XX Novel plants of grass family e.g. rice plants with increased cold-
 PT resistance obtained by reinforcing the expression of gene encoding enzyme
 PT e.g. ascorbate peroxidase, involved in elimination of active oxygen.
 XX
 PS Claim 5; SEQ ID NO 1; 27pp; Japanese.
 XX
 CC This invention relates to novel plants of the grass family whose cold-
 CC resistance property is reinforced by reinforcing the expression of a gene
 CC encoding an enzyme involved in elimination of active oxygen. The plant is
 CC preferably rice and the gene is ascorbate peroxidase gene which is
 CC derived from Chlamydomonas. The ascorbate peroxidase gene comprises a
 CC chloroplast transfer sequence, which allows the gene product to be
 CC translocated to chloroplast. The ascorbate peroxidase gene also comprises
 CC an added thylakoid membrane proximity transfer sequence, which allows the
 CC gene product to translocate in the vicinity of chloroplast. The invention
 CC is useful as a food crop suitable for cultivation in cold regions. The
 CC plants have increased cold-resistance. The present sequence is that of
 CC the Chlamydomonas ascorbate peroxidase gene which is used during the
 CC creation of the plants of the invention.
 XX
 SQ Sequence 1376 BP; 224 A; 473 C; 469 G; 210 T; 0 U; 0 Other;

Query Match 70.4%; Score 17.6; DB 12; Length 1376;
 Best Local Similarity 83.3%; Pred. No. 2.1e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAGGCAAGTAGTCCACACCTATG 24
 |||||
 DB 1273 GCGGCAACAGTAGCCACACCTATG 1296
 |||||

Search completed: December 14, 2004, 22:43:57
 Job time : 13.7084 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:27:19 ; Search time 2.10302 Seconds
(without alignments)
8449.603 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25
Sequence: 1 gaggcagtagtgcacacactatg 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.4	69.6	350	4 US-09-513-999C-12278	Sequence 12278, A
2	17.2	68.8	381	4 US-09-513-999C-13199	Sequence 13199, A
3	16.8	67.2	62804	3 US-09-800-960-3	Sequence 3, Appl1
4	16.8	67.2	62804	4 US-10-086-960-3	Sequence 3, Appl1
5	16.6	66.4	795	4 US-09-270-767-1106	Sequence 1106, Ap
6	16.6	66.4	795	4 US-09-270-767-16388	Sequence 16388, A
7	16.6	66.4	2379	4 US-09-205-258-175	Sequence 175, App
8	16.2	64.8	46	1 US-08-171-389-311	Sequence 311, App
9	16.2	64.8	46	1 US-08-123-836-311	Sequence 311, App
10	16.2	64.8	46	2 US-08-475-228A-311	Sequence 311, App
11	16.2	64.8	46	3 US-08-482-080A-311	Sequence 311, App
12	16.2	64.8	46	3 US-09-354-947-311	Sequence 311, App
13	16.2	64.8	46	5 PCT-US93-12388-311	Sequence 311, App
14	16.2	64.8	50	4 US-08-956-171B-4338	Sequence 4338, Ap
15	16.2	64.8	50	4 US-08-956-171B-5107	Sequence 5107, Ap
16	16.2	64.8	223	4 US-08-781-986A-5107	Sequence 5107, Ap
17	16.2	64.8	223	4 US-09-032-297A-12	Sequence 12, Appl
18	16.2	64.8	223	4 US-09-229-151C-6	Sequence 6, Appl1
19	16.2	64.8	335	4 US-08-956-171E-4200	Sequence 4200, Ap
20	16.2	64.8	335	4 US-08-781-986A-4200	Sequence 4200, Ap
21	16.2	64.8	339	4 US-08-956-171B-4338	Sequence 4338, Ap
22	16.2	64.8	339	4 US-08-781-986A-4338	Sequence 4338, Ap
23	16.2	64.8	463	4 US-08-956-171E-3850	Sequence 3850, Ap
24	16.2	64.8	463	4 US-08-781-986A-3850	Sequence 3850, Ap
25	16.2	64.8	467	4 US-09-270-767-14303	Sequence 14303, A
26	16.2	64.8	904	4 US-08-961-527-314	Sequence 314, App
27	16.2	64.8	1008	4 US-09-032-297A-7	Sequence 7, Appl1

28	16.2	64.8	1096	4 US-09-032-297A-2	Sequence 2, Appl1
29	16.2	64.8	1096	4 US-09-229-151C-2	Sequence 2, Appl1
30	16.2	64.8	1101	4 US-09-270-767-10150	Sequence 10150, A
31	16.2	64.8	1178	4 US-09-032-297A-1	Sequence 1, Appl1
32	16.2	64.8	1178	4 US-09-229-151C-1	Sequence 1, Appl1
33	16.2	64.8	1179	4 US-09-248-796A-2696	Sequence 2696, Ap
34	16.2	64.8	1357	3 US-09-448-176E-9	Sequence 9, Appl1
35	16.2	64.8	1357	3 US-09-448-176E-10	Sequence 10, Appl
36	16.2	64.8	1357	3 US-09-448-176E-11	Sequence 11, Appl
37	16.2	64.8	1450	4 US-09-032-297A-5	Sequence 5, Appl1
38	16.2	64.8	1587	4 US-09-032-297A-8	Sequence 8, Appl1
39	16.2	64.8	1789	4 US-09-032-297A-6	Sequence 6, Appl1
40	16.2	64.8	1894	4 US-09-032-297A-9	Sequence 9, Appl1
41	16.2	64.8	2270	4 US-09-229-151C-13	Sequence 13, Appl
42	16.2	64.8	2570	4 US-09-229-151C-14	Sequence 14, Appl
43	16.2	64.8	3095	3 US-08-434-000A-7	Sequence 7, Appl1
44	16.2	64.8	3095	3 US-09-312-157-7	Sequence 7, Appl1
45	16.2	64.8	3634	3 US-09-166-186-1	Sequence 1, Appl1

ALIGNMENTS

```

RESULT 1
US-09-513-999C-12278/c
; Sequence 12278, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT NO. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 12278
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-12278

Query Match      69.6%; Score 17.4; DB 4; Length 350;
Best Local Similarity 94.7%; Pred. No. 22;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GAGCAGTAGTTCACAC 19
Db      173 GAGCAGTAGTTCACAC 155

RESULT 2
US-09-513-999C-13199
; Sequence 13199, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT NO. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13199

```

```
/ LENGTH: 381
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 318
/ OTHER INFORMATION: k=g or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 319
/ OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-13199

Query Match          68.8%; Score 17.2; DB 4; Length 381;
Best Local Similarity 86.4%; Pred. No. 28;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GGCAGTAGTGGCCACACTATG 24
DB 280 GGCAGAGTTTCCACACCTATG 301

RESULT 3
US-09-800-960-3/c
/ Sequence 3, Application US/09800960
/ Patent No. 6387677
/ GENERAL INFORMATION:
/ APPLICANT: YE, Jane et al.
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL001158
/ CURRENT APPLICATION NUMBER: US/09/800,960
/ CURRENT FILING DATE: 2001-03-08
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 62804
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(62804)
/ OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

Query Match          67.2%; Score 16.8; DB 3; Length 62804;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCAGTAGTGGCCACACT 21
DB 43074 AGCAGTAGTGGCCACACT 43055

RESULT 4
US-10-096-960-3/c
/ Sequence 3, Application US/10096960
/ Patent No. 6664085
/ GENERAL INFORMATION:
/ APPLICANT: YE, Jane et al.
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ FILE REFERENCE: CL001158DIV
/ CURRENT APPLICATION NUMBER: US/10/096,960
/ CURRENT FILING DATE: 2002-03-14
/ PRIOR APPLICATION NUMBER: 09/800,960
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 62804
```

```
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(62804)
/ OTHER INFORMATION: n = A,T,C or G
US-10-096-960-3

Query Match          67.2%; Score 16.8; DB 4; Length 62804;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCAGTAGTGGCCACACT 21
DB 43074 AGCAGTAGTGGCCACACT 43055

RESULT 5
US-09-270-767-1106
/ Sequence 1106, Application US/09270767
/ Patent No. 6703451
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1106
/ LENGTH: 795
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-1106

Query Match          66.4%; Score 16.6; DB 4; Length 795;
Best Local Similarity 82.6%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGCAGTAGTGGCCACACTATG 25
DB 629 GGCAGTAGCGCCACACTATG 651

RESULT 6
US-09-270-767-16388
/ Sequence 16388, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 16388
/ LENGTH: 795
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-16388

Query Match          66.4%; Score 16.6; DB 4; Length 795;
Best Local Similarity 82.6%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GGCAGTAGTGGCCACACTATG 25
DB 629 GGCAGTAGCGCCACACTATG 651

RESULT 7
US-09-205-258-175/c
```

```

; Sequence 175, Application US/09205258
; Patent No. 6525174
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898

```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 2379
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1881)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-205-258-175

```

```

Query Match          66.4%; Score 16.6; DB 4; Length 2379;
Best Local Similarity 76.0%; Pred. No. 85;
Matches 19; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy 1 GAGGCGAGTAGTTCACACACCTATG 25
Db 1699 GAGGCTAAGCTGCCACACCGAKG 1675

```

```

RESULT 8
; US-08-171-389-311
; Sequence 311, Application US/08171389
; Patent No. 5578444
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk B.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/171,389
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:

```

```

1      APPLICATION NUMBER:  US 07/996,783
2      FILING DATE:  23-DEC-1992
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER:  US 07/723,618
5      FILING DATE:  27-JUN-1991
6      PRIOR APPLICATION DATA:
7      APPLICATION NUMBER:  US
8      FILING DATE:  22-JUN-1993
9      ATTORNEY/AGENT INFORMATION:
10     NAME:  Fabian, Gary R.
11     REGISTRATION NUMBER:  33,875
12     REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
13     TELECOMMUNICATION INFORMATION:
14     TELEPHONE:  (415) 324-0880
15     TELEFAX:  (415) 324-0960
16     INFORMATION FOR SEQ ID NO:  311:
17     SEQUENCE CHARACTERISTICS:
18     LENGTH:  46 base pairs
19     TYPE:  nucleic acid
20     STRANDEDNESS:  double
21     TOPOLOGY:  linear
22     MOLECULE TYPE:  DNA (genomic)
23     HYPOTHETICAL:  NO
24     ORIGINAL SOURCE:
25     INDIVIDUAL ISOLATE:  Human gene for tumor necrosis factor
26     INDIVIDUAL ISOLATE:  (TNF-alpha)
27     US-08-171-389-311

```

Query Match	64.8%	Score 16.2;	DB 1;	Length 46;
Best Local Similarity	85.7%	Pred. No. 54;		
Matches 18; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Oy 2 AGGCAGTAGTTGCCACACCTA 22
 | | | | | | | | | |
 Db 26 AGGCAGTTGTTGGCACACACCA 46

RESULT 9
 US-08-123-936-311
 ; Sequence 311 Application US/08123936
 ; Patent No. 5726014
 GENERAL INFORMATION:
 APPLICANT: Edwards, Cynthia A.
 APPLICANT: Cantor, Charles R.
 APPLICANT: Andrews, Beth M.
 APPLICANT: Thurn, Lisa M.
 TITLE OF INVENTION: Screening Assay for the Detection of
 NUMBER OF INVENTIONS: DNA-Binding Molecules
 NUMBER OF SEQUENCES: 640
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Genelabs Technologies, Inc.
 STREET: 505 Penobscot Drive
 CITY: Redwood City
 STATE: CA
 COUNTRY: USA
 ZIP: 94063
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/123,936
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/996,783
 FILING DATE: 23-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/723,618
 FILING DATE: 27-JUN-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
US-08-123-936-311

Query Match	64.8%	Score 16.2;	DB 1;	Length 46;
Best Local Similarity	85.7%	Pred. No. 54;		
Matches 18; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	2	AGGAGTAGTGGCACACCTA	22
Db	26	AGCAGTTGTGGCACACCCA	46

RESULT 10 08475-228A-311
 US-08475-228A-311
 Sequence 311. Application US/08475228A
 Patent No. 5869241
 GENERAL INFORMATION:
 APPLICANT: Edward, Cynthia A.
 APPLICANT: Cantor, Charles R.
 APPLICANT: Andrews, Beth M.
 APPLICANT: Turin, Lisa M.
 APPLICANT: Fy, Kirk E.
 TITLE OF INVENTION: Sequence-Directed DNA Binding
 TITLE OF INVENTION: Molecules, Compositions and Methods
 NUMBER OF SEQUENCES: 664
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genelabs Technologies, Inc.
 STREET: 505 Penobscot Drive
 CITY: Redwood City
 STATE: CA
 COUNTRY: USA
 ZIP: 94063
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/475,228A
 FILING DATE: 06-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/123,936
 FILING DATE: 17-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/996,783
 FILING DATE: 23-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/723,618
 FILING DATE: 27-JUN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/081,070
 FILING DATE: 22-JUN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Stratford, Carol A.
 REGISTRATION NUMBER: 34,444
 REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE: HUMAN
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
US-08-475-228A-311

Query Match 64.8%; Score 16.2; DB 2; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGCAAGTGTGGCACACCTA 22
DB 26 AGGCAAGTGTGGCACACCA 46

RESULT 11
US-08-482-080A-311
Sequence 311, Application US/08482080A
Patent No. 6010849
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,080A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brady, John P.
REGISTRATION NUMBER: 39,118
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880
TELEFAX: (650) 324-0960

INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE: HUMAN
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
US-08-482-080A-311

Query Match 64.8%; Score 16.2; DB 3; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGCAAGTGTGGCACACCTA 22
DB 26 AGGCAAGTGTGGCACACCA 46

RESULT 12
US-09-354-947-311
Sequence 311, Application US/09354947
Patent No. 6384208
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/354,947
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/482,080
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/171,389
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brady, John P.
REGISTRATION NUMBER: 39,118
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 324-0880

TELEFAX: (650) 324-0960
INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
US-09-354-947-311

Query Match 64.8%; Score 16.2; DB 3; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGCAAGTGTGGCACACCTA 22
Db 26 AGGCAAGTGTGGCACACCCA 46

RESULT 13
PCT-US93-12388-311
Sequence 311, Application PC/TUS9312388
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
City: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12388
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 311:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human gene for tumor necrosis factor
INDIVIDUAL ISOLATE: (TNF-alpha)
PCT-US93-12388-311

Query Match 64.8%; Score 16.2; DB 5; Length 46;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 AGGCAAGTGTGGCACACCTA 22
Db 26 AGGCAAGTGTGGCACACCCA 46

RESULT 14
US-08-956-171E-5107
Sequence 5107, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:

APPLICANT: Charles Kunesh

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 5107:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 5107:

US-08-956-171E-5107

Query Match 64.8%; Score 16.2; DB 4; Length 50;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GCAGTAGTGGCACACCTATG 24
Db 7 GCAGTAATGGCACACATCTGTG 27

RESULT 15
US-08-781-986A-5107
Sequence 5107, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:

APPLICANT: Charles Kunach
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5235
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5107:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-5107

Query Match 64.8%; Score 16.2; DB 4; Length 50;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 GCAGTAGTGCACACCTATG 24
Db 7 GCAGTAATTCACACATCTGTG 27

Search completed: December 15, 2004, 02:56:26
Job time : 3.10302 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2004, 22:25:49 ; Search time 12.7717 Seconds
(without alignments)
10792.424 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25
Sequence: 1 gagcagctagctgcacacactatg 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US06_PUB_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	9	US-09-995-225-42
2	25	100.0	25	10	US-09-995-225-42
3	25	100.0	1202	14	US-10-094-417-1
4	25	100.0	1343	10	US-09-813-432-21
5	25	100.0	1343	15	US-10-174-364-21
6	25	100.0	1343	16	US-10-246-583-21
7	25	100.0	1343	17	US-10-689-832-21
8	25	100.0	1526	14	US-10-012-140-10
9	25	100.0	1826	16	US-10-450-590-6
10	25	100.0	2117	17	US-10-779-104-1
11	19.2	76.8	688	16	US-10-424-599-59902
12	18.2	72.8	1886	17	US-10-437-963-65084

13	18.2	72.8	147309	9	US-09-742-312-3	Sequence 3, Appl1
14	18.2	72.8	147309	15	US-10-436-185-3	Sequence 3, Appl1
15	17.8	71.2	275	18	US-10-425-115-19752	Sequence 19751, A
16	17.8	71.2	787	17	US-10-767-701-14254	Sequence 14254, A
17	17.6	70.4	519	13	US-10-027-632-94559	Sequence 94559, A
18	17.6	70.4	519	13	US-10-027-632-94560	Sequence 94560, A
19	17.6	70.4	519	13	US-10-027-632-305465	Sequence 305465, A
20	17.6	70.4	519	13	US-10-027-632-305466	Sequence 305466, A
21	17.6	70.4	519	15	US-10-027-632-94559	Sequence 94559, A
22	17.6	70.4	519	15	US-10-027-632-94560	Sequence 94560, A
23	17.6	70.4	519	15	US-10-027-632-305465	Sequence 305465, A
24	17.6	70.4	519	15	US-10-027-632-305466	Sequence 305466, A
25	17.6	70.4	542	16	US-10-424-599-128896	Sequence 128896, A
26	17.6	70.4	152501	17	US-10-116-231-4	Sequence 4, Appl1
27	17.4	69.6	208	10	US-09-814-353-339	Sequence 339, App
28	17.4	69.6	208	10	US-09-814-353-6722	Sequence 6722, App
29	17.4	69.6	248	10	US-09-814-353-19298	Sequence 19298, A
30	17.4	69.6	527	10	US-09-814-353-13107	Sequence 13107, A
31	17.4	69.6	674	9	US-09-764-870-272	Sequence 272, App
32	17.4	69.6	674	14	US-10-125-540-272	Sequence 272, App
33	17.4	69.6	1596	15	US-10-371-264-2	Sequence 2, Appl1
34	17.4	69.6	1596	15	US-10-371-099-2	Sequence 2, Appl1
35	17.4	69.6	1596	16	US-10-371-122-2	Sequence 2, Appl1
36	17.4	69.6	1596	16	US-10-373-567-2	Sequence 2, Appl1
37	17.4	69.6	1596	16	US-10-628-088-2	Sequence 2, Appl1
38	17.4	69.6	1666	15	US-10-371-264-3	Sequence 3, Appl1
39	17.4	69.6	1666	15	US-10-371-099-3	Sequence 3, Appl1
40	17.4	69.6	1666	16	US-10-371-122-3	Sequence 3, Appl1
41	17.4	69.6	1666	16	US-10-373-567-3	Sequence 3, Appl1
42	17.4	69.6	1666	16	US-10-628-088-3	Sequence 3, Appl1
43	17.4	69.6	1860	15	US-10-371-264-5	Sequence 5, Appl1
44	17.4	69.6	1860	15	US-10-371-099-5	Sequence 5, Appl1
45	17.4	69.6	1860	16	US-10-371-122-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-09-995-225-42
Sequence 42, Application US/09995225
Publication No. US20020193584A1
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huong T.
APPLICANT: Lowitz, Kevin P.
TITLE OR INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
TITLE OR INVENTION: Receptors
FILE REFERENCE: AREN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06

```

; PRIOR APPLICATION NUMBER: 60/230,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020193584a1e1 Sequence
US-09-995-225-42
```

```
Query Match          100.0%; Score 25; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GAGGCGAGTAGTGGCCACACCTATGG 25
        |||
Db       1 GAGGCGAGTAGTGGCCACACCTATGG 25
```

```

RESULT 2
US-09-995-225-42
; Sequence 42, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A3-Endogenous Versions of Human G
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-42
```

Query Match 100.0%; Score 25; DB 10; Length 25;

Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 GAGGCGAGTAGTGGCCACACCTATGG 25
        |||
Db       1 GAGGCGAGTAGTGGCCACACCTATGG 25
```

```

RESULT 3
US-10-094-417-1/c
; Sequence 1, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiaqiang
; APPLICANT: Cutler, Gene
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030045685A1e1 Receptors
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
; NAME/KEY: CDS
; LOCATION: (68)..(1129)
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
US-10-094-417-1
```

Query Match 100.0%; Score 25; DB 14; Length 1202;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 GAGGCGAGTAGTGGCCACACCTATGG 25
        |||
Db       1156 GAGGCGAGTAGTGGCCACACCTATGG 1132
```

```

RESULT 4
US-09-813-432-21/c
; Sequence 21, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Majmuder, Kamud
; APPLICANT: Spaderna, Steven K
; APPLICANT: Smithson, Glenda
; APPLICANT: Wezes, Peter S
; APPLICANT: Vermet, Corine A. M.
; TITLE OF INVENTION: No. US20030148485A1e1 Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729
; CURRENT APPLICATION NUMBER: US/09/813,432
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
```

PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/193,843
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
US-09-813-432-21

Query Match 100.0%; Score 25; DB 10; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTTGCCACCTATGG 25
Db 1090 GAGGCACTAGTTGCCACCTATGG 1066

RESULT 5

US-10-174-364-21/c
Sequence 21, Application US/10174364
Publication No. US20030216308A1
GENERAL INFORMATION:
APPLICANT: Anderson et al.
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-729CIP2
CURRENT APPLICATION NUMBER: US/10/174,364
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,836
PRIOR FILING DATE: 2000-03-29
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1060)
US-10-174-364-21

Query Match 100.0%; Score 25; DB 15; Length 1343;

Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTTGCCACCTATGG 25
Db 1090 GAGGCACTAGTTGCCACCTATGG 1066

RESULT 6

US-10-246-583-21/c
Sequence 21, Application US/10246583
Publication No. US2004005862A1
GENERAL INFORMATION:
APPLICANT: Majumder
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-729CIP2CON1
CURRENT APPLICATION NUMBER: US/10/246,583
CURRENT FILING DATE: 2002-12-06
PRIOR APPLICATION NUMBER: 10/174,364
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,768
PRIOR FILING DATE: 2000-03-20
PRIOR APPLICATION NUMBER: 60/190,972
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,199
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/191,947
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,665
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,657
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,984
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: 60/192,664
PRIOR FILING DATE: 2000-03-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 1343
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1060)
US-10-246-583-21

Query Match 100.0%; Score 25; DB 16; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGGCACTAGTTGCCACCTATGG 25
Db 1090 GAGGCACTAGTTGCCACCTATGG 1066

RESULT 7

US-10-689-832-21/c
Sequence 21, Application US/10689832
Publication No. US20040121380A1
GENERAL INFORMATION:
APPLICANT: Majumder, Kamud
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND AMINO ACIDS ENCODING SAME
FILE REFERENCE: 15966-729DIV1
CURRENT APPLICATION NUMBER: US/10/689,832
CURRENT FILING DATE: 2003-10-20
PRIOR APPLICATION NUMBER: 09/813,432
PRIOR FILING DATE: 2001-03-20
PRIOR APPLICATION NUMBER: 60/190,835
PRIOR FILING DATE: 2000-03-20

```
/ PRIOR APPLICATION NUMBER: 60/190,768
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: 60/190,972
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,199
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: 60/191,947
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,665
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,657
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,984
/ PRIOR FILING DATE: 2000-03-28
/ PRIOR APPLICATION NUMBER: 60/192,664
/ PRIOR FILING DATE: 2000-03-28
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 21
/ LENGTH: 1343
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-689-832-21
```

```
Query Match      100.0%; Score 25; DB 17; Length 1343;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GAGGCACTAGTGGCCACCTATGG 25
Db      1090 GAGGCACTAGTGGCCACCTATGG 1066
```

```
RESULT 8
US-10-012-140-10/c
/ Sequence 10, Application US/10012140
/ Publication No. US2003009017A1
/ GENERAL INFORMATION:
/ APPLICANT: Leiby, Kevin R.
/ APPLICANT: Kappeler-Liebermann, Rosana
/ TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
/ TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
/ TITLE OF INVENTION: THEREFOR
/ FILE REFERENCE: 381552004900
/ CURRENT APPLICATION NUMBER: US/10/012,140
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: 60/246,768
/ PRIOR FILING DATE: 2000-11-08
/ PRIOR APPLICATION NUMBER: 60/246,772
/ PRIOR FILING DATE: 2000-11-08
/ PRIOR APPLICATION NUMBER: 60/249,185
/ PRIOR FILING DATE: 2000-11-15
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10
/ LENGTH: 1526
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (139)...(1200)
US-10-012-140-10
```

```
Query Match      100.0%; Score 25; DB 14; Length 1526;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GAGGCACTAGTGGCCACCTATGG 25
Db      1227 GAGGCACTAGTGGCCACCTATGG 1203
```

```
RESULT 9
US-10-450-590-6/c
/ Sequence 6, Application US/10450590
/ Publication No. US20040076985A1
/ GENERAL INFORMATION:
/ APPLICANT: Bayer AG
/ TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
/ FILE REFERENCE: L10316 Foreign Countries
/ CURRENT APPLICATION NUMBER: US/10/450,590
/ CURRENT FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/254,923
/ PRIOR FILING DATE: 2000-12-14
/ PRIOR APPLICATION NUMBER: US 60/280,110
/ PRIOR FILING DATE: 2001-04-02
/ PRIOR APPLICATION NUMBER: US 60/299,474
/ PRIOR FILING DATE: 2001-06-21
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 6
/ LENGTH: 1826
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-450-590-6
```

```
Query Match      100.0%; Score 25; DB 16; Length 1826;
Best Local Similarity 100.0%; Pred. No. 0.063;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GAGGCACTAGTGGCCACCTATGG 25
Db      1132 GAGGCACTAGTGGCCACCTATGG 1108
```

```
RESULT 10
US-10-779-104-1/c
/ Sequence 1, Application US/10779104
/ Publication No. US20040161799A1
/ GENERAL INFORMATION:
/ APPLICANT: Andrew J. Murphy
/ APPLICANT: Susan Croll-Kalish
/ TITLE OF INVENTION: KOR31like-Proteins and Methods of Modulating KOR3L-Mediated Activi
/ FILE REFERENCE: REG 1000A
/ CURRENT APPLICATION NUMBER: US/10/779,104
/ CURRENT FILING DATE: 2004-02-13
/ PRIOR APPLICATION NUMBER: 60/447,447
/ PRIOR FILING DATE: 2003-02-14
/ PRIOR APPLICATION NUMBER: 60/495,577
/ PRIOR FILING DATE: 2003-08-14
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 2117
/ TYPE: DNA
/ ORGANISM: homo sapiens
US-10-779-104-1
```

```
Query Match      100.0%; Score 25; DB 17; Length 2117;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GAGGCACTAGTGGCCACCTATGG 25
Db      1368 GAGGCACTAGTGGCCACCTATGG 1344
```

```
RESULT 11
US-10-424-599-59902/c
/ Sequence 59902, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
```

```

; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 59902
; LENGTH: 688
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_25101C.1
US-10-424-599-59902

Query Match          76.8%; Score 19.2; DB 16; Length 688;
Best Local Similarity 87.5%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGCAGTAGTGGCCACACCTATG 24
DB      353 GAGCAGTAGTGGCCGACCAATG 330

RESULT 12
US-10-437-963-65084/C
; Sequence 65084, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 65084
; LENGTH: 1886
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_66167C.1
US-10-437-963-65084

Query Match          72.8%; Score 18.2; DB 17; Length 1886;
Best Local Similarity 87.0%; Pred. No. 1,2e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGCAGTAGTGGCCACACCTAT 23
DB      78 GAGCAGTAGTGGCCATAGCTAT 56

RESULT 13
US-09-742-312-3
; Sequence 3, Application US/09742312
; Patent No. US20020045166A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL000838
; CURRENT APPLICATION NUMBER: US/09/742,312
; CURRENT FILING DATE: 2000-12-22
```

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 147309
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(147309)
; OTHER INFORMATION: n = A,T,C or G
US-09-742-312-3

Query Match          72.8%; Score 18.2; DB 9; Length 147309;
Best Local Similarity 87.0%; Pred. No. 2,1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGCAGTAGTGGCCACACCTAT 23
DB      122948 GAGCAGTAGTGGCCATTCTAT 122970

RESULT 14
US-10-436-185-3
; Sequence 3, Application US/10436185
; Publication No. US20030180887A1
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL000838CON
; CURRENT APPLICATION NUMBER: US/10/436,185
; CURRENT FILING DATE: 2003-05-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 147309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(147309)
; OTHER INFORMATION: n = A,T,C or G
US-10-436-185-3

Query Match          72.8%; Score 18.2; DB 15; Length 147309;
Best Local Similarity 87.0%; Pred. No. 2,1e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GAGCAGTAGTGGCCACACCTAT 23
DB      122948 GAGCAGTAGTGGCCATTCTAT 122970

RESULT 15
US-10-425-115-19752
; Sequence 19752, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 19752
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Zea mays
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_118018C.1
US-10-425-115-19752

Query Match      71.2%; Score 17.8; DB 18; Length 275;
Best Local Similarity 90.5%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GAGGCAGTAGTGGCCACGCT 21
        ||| ||||| ||||| |||||
Db      103 GAGCGATGATGCCACGCT 123
```

Search completed: December 15, 2004, 02:53:13
Job time : 13.7717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: December 14, 2004, 22:17:59 ; Search time 78.7335 Seconds
(without alignments)
11570.599 Million cell updates/sec

Title: US-09-995-225B-42

Perfect score: 25

Sequence: 1 gagcagtagtcgcacacactatg 25

Scoring table: IDENTITY_NUC

Searched: 3282875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : Listing first 45 summaries

EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	448	8	AQ225693 HS_2009_B
2	20.2	80.8	746	6	CD636851 56003360H
3	19	76.0	590	9	CL686359 PRI0144A
4	19	76.0	808	2	CL692939 PRI015B_F
5	18.8	75.2	172	2	AW799209 RCO-UM005
6	18.8	75.2	334	1	AI029880 UT-R-CO-1
7	18.8	75.2	565	6	CD636852 56003360J
8	18.8	75.2	697	5	BU616836
9	18.8	75.2	729	7	CO138524 EST833195
10	18.6	74.4	453	8	AZ009690 RPCI-23-3
11	18.6	74.4	483	4	B2260234
12	18.6	74.4	489	4	BI680378
13	18.6	74.4	495	4	BI260587
14	18.6	74.4	504	5	BI260587
15	18.6	74.4	506	2	AX499223 DKF2P779E
16	18.6	74.4	543	1	AW591426 AV591426
17	18.6	74.4	546	2	BF076557 226150 MA
18	18.6	74.4	557	4	BG691654 340911 BA
19	18.6	74.4	572	4	BM030860 495388 MA
20	18.6	74.4	590	2	BE590190
21	18.6	74.4	590	4	BI682340 463396 MA
22	18.6	74.4	613	1	AL781067
23	18.6	74.4	619	1	AL802528
24	18.6	74.4	621	4	BU260398

25	18.6	74.4	639	1	AL656915
26	18.6	74.4	640	1	AL876764
27	18.6	74.4	642	1	AL877073
28	18.6	74.4	652	7	CK835565
29	18.6	74.4	661	1	AL629682
30	18.6	74.4	708	7	CK951479
31	18.6	74.4	745	7	CK832878
32	18.6	74.4	774	5	BU228619
33	18.6	74.4	785	5	EX739364
34	18.6	74.4	803	9	CC905917
35	18.4	73.6	294	6	CD850431
36	18.4	73.6	311	6	CD850445
37	18.4	73.6	407	6	CB807057
38	18.4	73.6	519	6	CD849483
39	18.4	73.6	591	6	CD849077
40	18.4	73.6	626	6	CD849065
41	18.4	73.6	670	6	CD848640
42	18.4	73.6	675	6	CD848811
43	18.4	73.6	682	6	CD850036
44	18.4	73.6	683	6	CD850069
45	18.4	73.6	689	6	CD849410

ALIGNMENTS

RESULT 1
AQ225693 448 bp DNA linear GSS 26-SEP-1998
LOCUS
DEFINITION HS_2009_B2_B08_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2009 Col=16 Row=D, genomic survey

ACCESSION AQ225693
VERSION AQ225693.1 GI:3650922
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mahairas,G.C., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764

COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2009 row: D column: 16
Class: BAC ends
High quality sequence stop: 448.
Location/Qualifiers

FEATURES
source

1. 448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=2009 Col=16 Row=D"
/sex="male"
/clone_1b="CIT Approved Human Genomic Sperm Library D"
/note="Organ: Sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

ORIGIN
Query Match 100.0%; Score 25; DB 8; Length 448;
Best Local Similarity 100.0%; Pred. No. 0.66;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCAGTAGTTCACACCTATGG 25
 |||||
 151 GAGGCAGTAGTTCACACCTATGG 175

RESULT 2
 LOCUS CD636851/c 746 bp mRNA linear EST 12-JUN-2004
 DEFINITION S6003360H1 FLP Homo sapiens cDNA, mRNA sequence.
 ACCESSION CD636851
 VERSION CD636851.1 GI:40285118
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 746)
 Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
 TITLE Circular rapid amplification of cDNA ends for high-throughput
 extension cloning of partial genes
 Genomics 84 (1), 205-210 (2004)
 CONTACT Contact: Fu GK
 INCYTE Genomics, Inc.
 3160 Porter Dr., Palo Alto, CA 94304, USA
 Tel: 6508454102
 Email: gfu@incyte.com.

FEATURES
 Source 1..746
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="FLP"
 /note="Vector: pDrive Cloning Vector"

ORIGIN
 Query Match 80.8%; Score 20.2; DB 6; Length 746;
 Best Local Similarity 88.0%; Pred. No. 1.3e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGGCAGTAGTTCACACCTATGG 25
 |||||
 658 GTGGCAGAGTTCCACACTATGG 634

RESULT 3
 LOCUS CL686359 590 bp DNA linear GSS 09-JUN-2004
 DEFINITION PRI0144a_B01_2 - PRI0144a.BR (590) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.
 ACCESSION CL686359
 VERSION CL686359.1 GI:50194822
 KEYWORDS GSS.
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 590)
 Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
 TITLE AppADB: an Acedb database for the nematode satellite organism
 Pristionchus pacificus
 Nucleic Acids Res. 32 (1), D421-D422 (2004)
 CONTACT Contact: Sommer RJ
 JOURNAL Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.

Seq primer: T7
 Class: fosmid ends.
 Location/Qualifiers
 1..590
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
 Query Match 76.0%; Score 19; DB 9; Length 590;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCAGTAGTTCACAC 19
 |||||
 528 GAGGCAGTAGTTCACAC 546

RESULT 4
 LOCUS CL692939 808 bp DNA linear GSS 10-JUN-2004
 DEFINITION PRI015b_F10_2 - PRI015b.BR (808) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.
 ACCESSION CL692939
 VERSION CL692939.1 GI:50214847
 KEYWORDS GSS.
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 808)
 Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
 TITLE AppADB: an Acedb database for the nematode satellite organism
 Pristionchus pacificus
 Nucleic Acids Res. 32 (1), D421-D422 (2004)
 CONTACT Contact: Sommer RJ
 JOURNAL Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: T7
 Class: fosmid ends.
 Location/Qualifiers
 1..808
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
 Query Match 76.0%; Score 19; DB 9; Length 808;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGCAGTAGTTCACAC 19
 |||||
 528 GAGGCAGTAGTTCACAC 546

RESULT 5
 AW799209

[illegible]

```

REFERENCE      1 (bases 1 to 334)
AUTHORS       Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE         Normalization and subtraction: two approaches to facilitate gene
              discovery
JOURNAL       Genome Res. 6 (9), 791-806 (1996)
MEDLINE       97044477
PUBMED        8889548
COMMENT       Contact: Soares, MB
              Coordinated Laboratory for Computational Genomics
              University of Iowa
              375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
              Tel: 319 335 8250
              Fax: 319 335 9565
              Email: bento-soares@iowa.edu
              Oligo-dt track not found, Not 1 site shown in beginning of sequence
              is likely internal to the message. cDNA Library Preparation: M.
              Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
              through Research Genetics This clone is also available through the
              I.M.A.G.E. Consortium at LNL (info@image.llnl.gov). IMAGE
              ID=1783323
              Seq primer: M13 Forward
              POLYA=No.

FEATURES
SOURCE
Location/Qualifiers
1..334
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CO-1f-d-11-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_id="UI-R-CO"
/note="Vector: pTZ19D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CO
library is a subcloned library derived from the UI-R-A1
and UI-R-E1 libraries. The UI-R-A1 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, and muscle. The UI-R-E1
library consisted of a mixture of individually tagged
normalized libraries constructed from 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dt track which
allows identification of the library of origin of a clone
within the mixture. The subtracted library (UI-R-CO) was
constructed as follows: PCR amplified cDNA inserts from a
pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the pooled UI-R-A1 and UI-R-E1 library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-CO
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)"

```

```

ACCESSION   CD636852.1 GI:40285119
VERSION     EST.
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 565)
AUTHORS     Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE       Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
JOURNAL     Genomics 84 (1), 205-210 (2004)
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.

FEATURES
  source
    1..565
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone_lib="RIP"
    /note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match      75.2%; Score 18.8; DB 6; Length 565;
Best Local Similarity 90.9%; Pred. No. 5.8e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      3 GGCAGTGTGGCCACCTATG 24
        |||||
Db      170 GGCAGAGTTGCCACCTATG 191

RESULT 8
BU616836      697 bp mRNA linear EST 23-SEP-2002
DEFINITION   UT-H-FH1-bfj-c-18-0-UI.s1 NCI CGAP FH1 Homo sapiens cDNA clone
LOCUS        BU616836
ACCESSION    BU616836
VERSION      BU616836.1 GI:23283051
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 697)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL      Tumor Gene Index
COMMENT      Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapsb-remail.nih.gov
            Tissue Procurement: James Martin
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=yes.

FEATURES
  source
    1..697
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone_lib="UI-H-FH1-bfj-c-18-0-UI"
    /tissue_type="Cell Line"
    /dev_stage="Adult"
    /lab_host="DH10B (Life Technologies)"
    /clone_lib="NCI CGAP FH1"
    /note="Organ: Chondrosarcoma; Vector: pTT73-Pac

```

```

(Pharmacia) with a modified polylinker. Site 1: EcoR I;
Site 2: Not I; NCI CGAP_FH1 is a normalized cDNA library
obtained from a cell line derived from grade I
chondrosarcoma tissue. The library was constructed and
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pTT73-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AGATCCGGC. The cell line was provided by Dr. James Martin
from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Cell Line C58 - Grade 1
Chondrosarcoma
TAG_LIB=UI-H-FH1
TAG_SEQ=AGATCCGGC"

ORIGIN
Query Match      75.2%; Score 18.8; DB 5; Length 697;
Best Local Similarity 90.9%; Pred. No. 6e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAGGACGAGTGTGCCACCTA 22
        |||||
Db      288 GAGGACGAGTGTGCCACCTA 309

RESULT 9
COL18524      729 bp mRNA linear EST 17-JUN-2004
LOCUS        COL18524
DEFINITION   EST833195 Aspergillus flavus Normalized cDNA Expression Library
ASpergillus flavus clone NAFU615' end similar to (Q7PG55)
ENSANGP0000023647 (Fragment), mRNA sequence.
ACCESSION    COL18524
VERSION      COL18524.1 GI:48887502
KEYWORDS     EST.
SOURCE       Aspergillus flavus
ORGANISM     Aspergillus flavus
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillu.
REFERENCE    1 (bases 1 to 729)
AUTHORS      Yu,J., Whitelaw,C.A., Nierman,W.C., Bhatnagar,D. and Cleveland,T.E.
TITLE        genes with putative roles in aflatoxin contamination of crops
JOURNAL      FEMS Microbiol. Lett. (2004) In press
COMMENT      Contact: Yu J
            Food and Feed Safety Research Unit
            USDA/ARS, Southern Regional Research Center
            1100 Robert R. Lee Boulevard, New Orleans, LA 70124, USA
            Tel: 504 286 4405
            Fax: 504 286 4419
            Email: jiyu@ars.ars.usda.gov
            Contact Dr. Yu at USDA/ARS SRRC (jiyu@ars.ars.usda.gov) for clone
            information
            PCR primers
            FORWARD: M13P
            BACKWARD: M13R
            Seq primer: M13 Forward
            POLYA=no.

FEATURES
  source
    1..729
    /organism="Aspergillus flavus"
    /mol_type="mRNA"
    /strain="NRRL 3357"
    /db_xref="taxon:5059"
    /clone="NAFCU61"
    /sex="asexual mycelia"
    /cell_type="mycelia"
    /dev_stage="developmental stages from 18 to 96 hours"
    /lab_host="E. coli DH10B T1 resistant cells"

```

/clone_11b="Asepergillus flavus Normalized cDNA Expression Library"
/note="Vector: pBluescript (SK-) (Stratagene), antibiotic selection marker: Carbenicillin; Site_1: NotI, at the 5 prime end; Site_2: EcoRI, at the 3 prime end; This normalized cDNA expression library was constructed using a mixture of mycelial cells grown under eight different medium conditions and harvested at 5 time points (18, 24, 48, 72, 96 hours). The poly-A sequence was trimmed off before ligating to vector."

ORIGIN

Query Match 75.2%; Score 18.8; DB 7; Length 729;

Best Local Similarity 90.9%; Pred. No. 6e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 20; Conservative 0; Gaps 0; Indels 0; Gaps 0;

Qy 2 AGGCAGTAGTGGCCACACTAT 23

Db 652 AGGCATATGTTGCCACACCAT 673

RESULT 10 A2009690 453 bp DNA linear GSS 25-FEB-2000

LOCUS RPCI-23-352C3.TV RPCI-23 Mus musculus genomic clone RPCI-23-352C3,

DEFINITION genomic survey sequence.

ACCESSION A2009690

VERSION A2009690.1 GI:7085074

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 453)

Zhao, S., Nieman, W., Feldblum, T., Malek, J., Shatsman, S.,

Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de

Jong, P. and Fraser, C. M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1998)

Other GSSes: RPCI-23-352C3.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)

or from Reesha Ch Genetics (info@reesgen.com). BAC end page:

http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html

plate: 352 row: C column: 3

Seq primer: T7

Class: BAC ends

FEATURES

Source

Location/Qualifiers

1. 453

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-352C3"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1:

EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcoRI and EcoRV. Size

selected DNA was cloned into the pBAC3.6 vector at the

EcoRI sites. The ligation products were transformed into

ORIGIN

Query Match 74.4%; Score 18.6; DB 8; Length 453;

Best Local Similarity 84.0%; Pred. No. 7e+02; Mismatches 4; Indels 0; Gaps 0;

Matches 21; Conservative 0; Gaps 0; Indels 0; Gaps 0;

Qy 1 GAGGCAGTAGTGGCCACACTATG 25

Db 365 GAGGCATGCTTGGCCACACCATG 341

RESULT 11

Bj260234/c

LOCUS Bj260234

DEFINITION Bj260234 Y. Ogihara unpublished cDNA library, Wh_h Triticum

aestivum cDNA clone wh24a16 5', mRNA sequence.

ACCESSION Bj260234

VERSION Bj260234.1 GI:20082291

KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)

ORGANISM Triticum aestivum

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.

1 (bases 1 to 483)

Ogihara, Y. and Murai, K.

Expressed genes in Triticum aestivum

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genie.nig.ac.jp.

Location/Qualifiers

1. 483

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Chinese Spring"

/db_xref="taxon:4565"

/clone="wh24a16"

/issue_type="spike at heading date"

/dev_stage="Feekes' scale 10.5"

/clone_11b="Y. Ogihara unpublished cDNA library, Wh_h"

ORIGIN

Query Match

Best Local Similarity

Matches

Qy

Db

RESULT 12

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 489)

Smith, T. P. L., Grose, W. M., Freking, B. A., Roberts, A. J., Stone, R. T.,

Casas, E., Wray, J. E., White, J., Cho, J., Fahnenkrug, S. C.,

Bennett, G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A.,

Chitko-McKown, C. G., Perata, G., Holt, I., Karamycheva, S., Liang, F.,

TITLE Quackenbush, J. and Keele, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
PubMed 11282978

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smtlth@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.
PCR Primers

FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 126 row: C column: 19
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers

FEATURES

source

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/issue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1BOV"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

ORIGIN

Query Match 74.4%; Score 18.6; DB 4; Length 489;
Best Local Similarity 84.0%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAGGAGTAGTTCACACCTATGG 25
Db 144 GAGGAGCAGTTCACCTCAGTGG 120

RESULT 13 495 bp mRNA linear EST 08-APR-2002
LOCUS BJ260587/c
DEFINITION BJ260587 Y. Ogihara unpublished cDNA library, wh_h Triticum
aestivum cDNA clone whn25m17 5', mRNA sequence.
ACCESSION BJ260587 GI:20082510
VERSION BJ260587.1
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 495)
Ogihara, Y. and Murai, K.
Expressed genes in Triticum aestivum
Unpublished (2002)
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source

1..495
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whn25m17"

/issue_type="spike at heading date"
/dev_stage="Fleeker scale 10.5"
/clone_lib="Y. Ogihara unpublished cDNA library, wh_h"

ORIGIN

Query Match 74.4%; Score 18.6; DB 4; Length 495;
Best Local Similarity 84.0%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAGGAGTAGTTCACACCTATGG 25
Db 269 GAGGAGCAGTTCACCAACATATGG 245

RESULT 14 504 bp mRNA linear EST 04-SEP-2003
LOCUS BX499223/c
DEFINITION DKFZP779E0944.r1.779 (synonym: hnccl) Homo sapiens cDNA clone
DKFZP779E0944.5, mRNA sequence.
ACCESSION BX499223
VERSION BX499223.1 GI:32017467
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Bioecker, H., Boecker, M., Mewes, H.W., Well, B., Aml, C., Oanger, A.,
1 (bases 1 to 504)
Fob, G., Han, M., and Wiemann, S.
EST (Bioecker, H., Boecker, M., Mewes, H.W., Well, B., Aml, C., et al.)
Unpublished (2003)
Contact: MIPS

COMMENT

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GRF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZP779E0944) is available at the RZPD in Berlin.
Please contact the RZPD: Resourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP779E0944"
/issue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/clone_lib="779 (synonym: hnccl)"
/note="Vector: pSPORT1_Sfi; Site_1: SfiI; Site_2: SfiIB"

ORIGIN

Query Match 74.4%; Score 18.6; DB 5; Length 504;
Best Local Similarity 84.0%; Pred. No. 7.1e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAGGAGTAGTTCACACCTATGG 25
Db 496 GAGGAGCAGTTCACCTCAGTGG 472

RESULT 15 506 bp mRNA linear EST 25-APR-2001
LOCUS AM653346/c
DEFINITION 101999 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AM653346
VERSION AM653346.1 GI:7419172
KEYWORDS EST.

SOURCE
ORGANISM

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.REFERENCE
AUTHORS

1 (bases 1 to 506)

Smith, T.P.L., Grose, W.M., Preking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,
Chitko-Mckown, C.G., Peltas, G., Holt, I., Karamycheva, S., Liang, F.,
Quackenbush, J., and Keele, J.W.Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)JOURNAL
MEDLINE
PUBMED21180013
11282978

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

PCR primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCACGACG

Plate: 104 row: P column: 14

Seq primer: ATTGAGTGACACTATAG.

Location/Qualifiers

FEATURES
source

1..506

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/cissue_type="pooled"

/lab_host="DH10B"

/clone_id="MARC 1BOV"

/note="Vector: PCMV SPOT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

ORIGIN

Query Match

74.4%; Score 18.6; DB 2; Length 506;

Best Local Similarity

84.0%; Pred. No. 7.1e+02;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAGCAGTAGTGCACACCTATGG 25

DB 86 GAGCAGCACTTCCACTCCACGTGG 62

Search completed: December 15, 2004, 02:35:10
Job time : 81.7335 secs

This Page Blank (uspto)